High-Efficiency Microfluidics for Single Cell Phenotypic and Transcriptomic Analysis of Rare Cancer Cells

by

Yu-Heng Cheng

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical and Computer Engineering)
in The University of Michigan
2018

Doctoral Committee:

Professor Euisik Yoon, Chair
Dr. Yu-Chih Chen
Assistant Professor Somin E. Lee
Professor Max S. Wicha
Emeritus Professor Kensall D. Wise

Yu-Heng Cheng

yhcheng@umich.edu

ORCID iD: 0000-0002-8627-7160

DEDICATION

For Hao Chen-Hung and Liangwang Chen

ACKNOWLEDGEMENTS

I still remember the day I first learned about genetic codes in a biology class. I was so fascinated that I told myself this is something I want to further investigate later. It is truly a dream coming true for me to develop technologies to decipher cancer cell gene expressions and push our understanding of oncology further. This wonderful journey wouldn't have happened without the support of many friends and colleagues.

First, I would like to thank my Ph.D. research advisor, Professor Euisik Yoon, for the opportunity to join his lab and motivating me to pursue ambitious research goals. I have learned a lot during my time at Michigan, and my mindset and skills cultivated here will be invaluable in my future career.

I would also like to acknowledge my dissertation committee members, Dr. Yu-Chih Chen, Professor Somin Lee, Professor Max Wicha, and Professor Ken Wise. Their feedback and encouragement have helped me developed my research and dissertation significantly. I greatly appreciate the collaboration with Professor Max Wicha and Dr. Monika Burness. They have been providing valuable advice from oncologists' perspective and supporting my research with clinical samples for experiments. I am also very grateful for Dr. Yu-Chih Chen for his hands-on guidance and advice as we have been working closely for six years. I would also extend my appreciation to Prof. Hyun Min Kang, Prof.

Sunitha Nagrath, Dr. Eric Lin, Dr. Michael Brooks, Dr. Ramdane Harouaka, Anoop Gill, and Dr. Ebrahim Azizi for their help in my research.

Dr. Patrick Ingram, Dr. Seungwon Jung, Zhixiong Zhang, Riley Brien, and Woncheol Lee each deserve special recognition for their contributions to this research as well. Each of them has supported me considerably in brainstorming, in research, and through their friendship. I would finally thank all the current and past members of the Yoon lab.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

Cancer is one of the leading causes of death worldwide, with recent research suggesting that a high degree of cancer cellular heterogeneity leads to different responses to therapies, making cancer a challenging disease to cure. As such, characterizing tumor cells at single-cell resolution promises greater insight into the mechanisms of cancer progression and facilitates the development of cancer treatments targeting different cell types. In recent years, microfluidics has emerged as a promising platform technology for single cell analysis. However, most existing single-cell analysis platforms cannot be applied to study certain rare, yet important, cancer cell populations. Cancer stem-like cells (CSCs), for instance, represent a small subpopulation (1-5%) of the tumor cells, but they are tumorigenic and cause tumor relapse and metastasis, which is the cause of over 90% cancer related death. In addition, samples harvested from microfluidic assays represent another critical rare-cell population for study. As microfluidics has become ubiquitous in labs for cellular assays, there is an emerging need to interface other microfluidics to further investigate the cells of interest. Finally, liquid biopsy of circulating tumor cells (CTCs) analysis has a high potential for cancer diagnostics and precision medicine. However, CTCs are very rare in blood with the concentration of only 1-20 cells/mL, which is a great challenge for cell capture and the downstream analysis. To achieve single-cell analysis of these rare cell populations, this thesis presents high-efficiency microfluidic technologies for single cell phenotypic and transcriptomic studies.

First, to study the CSCs, we achieved **the scaling and automation of high-throughput single-cell-derived tumor sphere assays**, which are a strong indicator of disease outcome (e.g. tumor relapse and metastasis). With a highly parallel chamber capture structure, the array size of the chip can be scaled from 800 to 12,800 while maintaining a high single cell capture rate of ~76.5% to study the heterogeneity of CSCs. The assay elucidated a controversial hypothesis of the linkage between cell size and tumor-initiating potential. The cell capture scheme was also applied to cell-to-cell interaction and cell differentiation studies, highlighting its versatility in single cell analysis. Second, to interface the sample from other microfluidics, a **proteolytic chip with a vacuum-driven single-cell capture scheme** was developed to enable handling of small number of cells down to ~50 cells for a given sample volume of 4 μl. By a vacuum driven cell loading process, the entire input solution can be loaded into each cell capture chamber to minimize cell loss in the dead volume. After loading protease sensitive reagents through diffusion and performing air isolation between capture chambers, the proteolytic activity, an important process in metastasis, of each cell can be monitored at single cell resolution. Finally, **Hydro-Seq**, a scalable hydrodynamic bead-cell-pairing technique, was developed to analyze CTCs with high cell capture efficiency, high-throughput, and contamination removal capability. We successfully achieved whole transcriptome sequencing of 666 CTCs from 21 breast cancer patient samples, identifying critical cancer metastasis subpopulations of mesenchymal–epithelial transition (MET) and epithelial–mesenchymal transition (EMT) with a fraction of cells expressing epithelial and mesenchymal CSC markers such as CD44+/CD24- and ALDH. The presented technology offers the capability

to analyze the phenotype and transcriptome of rare cancer cells, ultimately providing better diagnostics and treatment of cancer in the future.

Chapter 1
INTRODUCTION

Cancer has a major impact on society worldwide, and the incidence rate has been increasing due to the prolonged human life expectancy and changes in lifestyle [1–3]. As such, tremendous effort has been made to develop treatment for cancer. In 1971, former U.S. President Richard Nixon declared "a war on cancer" by signing the National Cancer Act of 1971 to augment funding and resources for cancer research [4]. Until recently, programs such as "Precision Medicine Initiative" and "Cancer Moonshot" are still being launched to continue the development of cancer therapy [5,6]. However, because of the high degree of cellular heterogeneity in tumor, it has been challenging to identify the rare tumor initiating cell populations and develop effective treatments for cancer [7,8]. Therefore, single-cell analysis has emerged as a promising method to study the heterogeneity at single cell resolution, and various microfluidics technologies have been developed for cell manipulation and on-chip assays as micro-Total-Analysis-Systems (MicroTAS). However, due to the limited throughput and severe cell loss during the cell capture process in those platforms, it is still difficult to interface microfluidics with rare cell populations to enable new research and clinical applications. For instance, liquid biopsy to collect circulating tumor cells (CTCs) for analysis is a promising way for non-invasive cancer diagnostics, but there are only around 1-100 CTCs from each blood draw, so such sample cannot be handled by most microfluidic platforms [9,10]. To achieve rare cell sample processing, this thesis focuses on the development of microfluidic technologies

with high capture efficiency and high throughput for single-cell phenotypic and transcriptomic analysis. This chapter will discuss three kinds of important rare cell populations and outlines the following chapters to introduce the engineering innovations to overcome the technical challenges.

1.1 Cancer and cancer heterogeneity

Cancer is the leading cause of death in many developed and developing countries [11]. Because of the instability of genomics and epigenetics regulation in cancer, cancer cells are known for their cellular heterogeneity and it poses great challenges in cancer treatment [12,13]. Different models have been proposed to explain such tumor heterogeneity. The cancer Stem-like Cells (CSCs) model suggests that an only a rare subpopulation of the tumor cells is tumorigenic and capable of forming new tumors. During cell proliferation, the CSCs would perform self-renewal to maintain the CSC population and cell differentiation to generate progenitor cells and differentiated cells that make up the majority of the tumor. As such, there is a differentiation hierarchy similar to that of stem

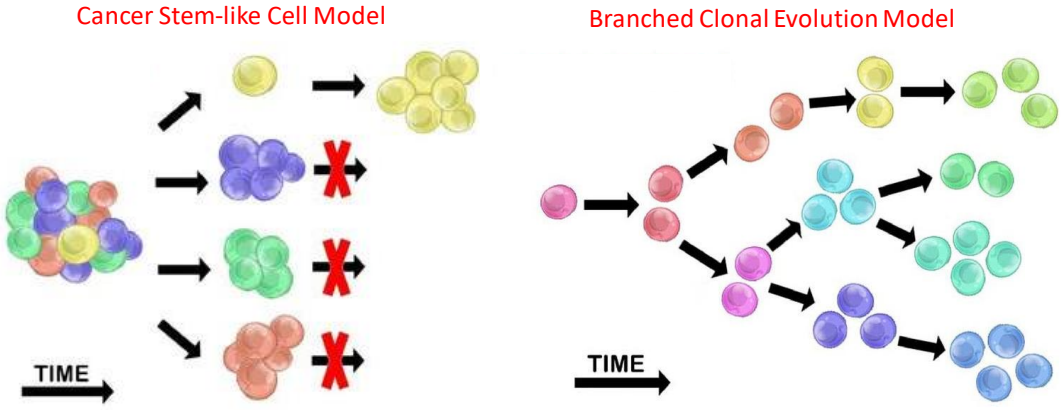

Figure 1-1 Two models to explain tumor heterogeneity (Left) The cancer stem-like cell (CSC) model: the rare CSC (yellow) subpopulation is capable of forming new tumors, while others cell populations are not tumorigenic. (Right) The branched clonal evolution model: the tumor cells accumulate random mutations and epimutations to give rise to some subpopulation evolutionary advantage overtime. [56][57]

cells to describe the tumor heterogeneity. To identify the CSCs and their biomarkers, researchers measure the tumorigenic potential of the tumor subpopulation by applying the xenograft model with injection of a limited number of tumor cells to immunodeficient mice. Concretely, when CSCs are implanted to the mice, they would grow into new tumors in the animals and thus are considered tumorigenic, while non-CSCs cannot form new tumors. Another model to explain the tumor heterogeneity is the clonal evolution model. This model suggests that the tumor cells accumulate random mutations and epimutations as they grow. This heterogeneity may give some subpopulation an evolutionary advantage in the tumor environment for tumor invasion and tumor relapse.

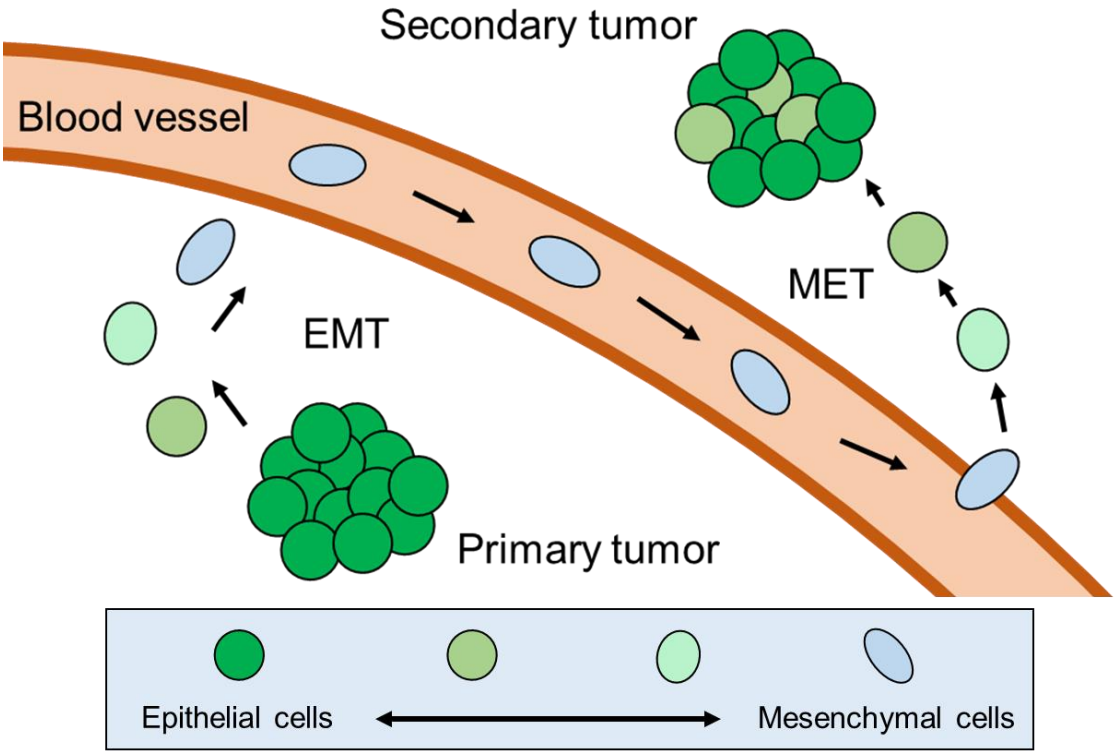

Figure 1-2 The epithelial to mesenchymal transition (EMT) and mesenchymal to epithelial transition (MET) during cancer metastasis

Tumor heterogeneity is also observed in the process of cancer metastasis, which is the major cause of cancer related death in many cancer diseases, as the tumor cells are found to perform transition between epithelial and mesenchymal cell types to overcome to

obstacles in metastasis[14,15]. In the primary tumor, most cells exhibit an epithelial cell property with an elevated proliferation rate and lower motility for cell migration. To metastasize to a distant location, some epithelial cells can perform the epithelial to mesenchymal transition (EMT) to acquire the mesenchymal cell property. The mesenchymal state cells have an increased cell motility to migrate and invade to the distant stroma, reaching the circulatory system such as lymph and blood vessels for intravasation. The mesenchymal phenotype also facilitates the cells to stay quiescent and dormant, so they can resist the programmed cell death due to the loss of cellular anchorage (termed anoikis) and survive the harsh environment during circulation. Finally, some metastasizing cells can reach to a distant tissue location by anchoring to the local tissue and exiting the circulatory system. To form the secondary tumor, the mesenchymal cells can perform the mesenchymal to epithelial transition (MET) to acquire the epithelial property. The epithelial cell state allows the cells to proliferate again for tumor growth and colonization. The cellular heterogeneity of EMT and MET facilitates the tumor cells to attain the required cell property for cancer metastasis.

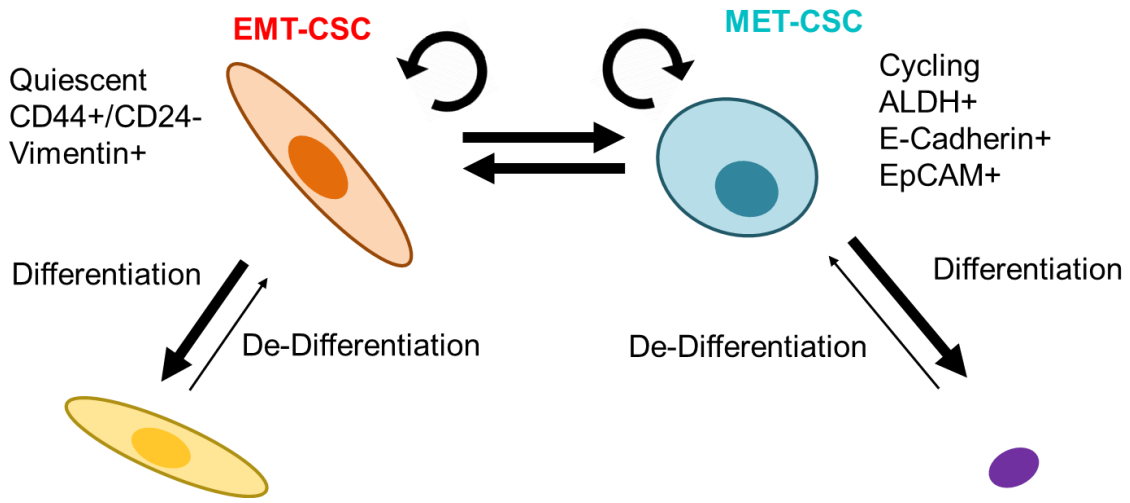

Figure 1-3 The cancer stem cell (CSC) hierarchy representing the CSCs with EMT and MET cell types.

To characterize those EMT and MET cells, different biomarkers have been developed to identify the cells [15,16]. The epithelial cell types are often identified by E-cadherin (CDH), Epithelial cell adhesion molecule (EpCAM), Keratin-8 (KRT8) and Keratin-18 (KRT18). Since the cells are cycling, cell cycle genes such as c-jun (JUN) and Cyclin D2 (CCND) may also be found in those epithelial cells. The mesenchymal cells are identified by the Plasminogen activator inhibitor-1 (SERPINE1), the structural protein Vimentin (VIN), Phosphoinositide-dependent kinase-1 (PDK1), the transcriptional factor Zinc finger E-box-binding homeobox 2 (ZEB2) and transforming growth factor beta (TGF-β).

The cancer stem cells are known for their cancer plasticity to perform the EMT and MET for metastasis [15,17,18]. In the cancer stem cell hierarchy, there are EMT-CSCs and MET-CSCs that can perform the transition between the epithelial and mesenchymal state. They are also able to perform self-renewal to maintain the CSC population. Additionally, they can also further differentiate into the regular cancer cells for tumor growth and expansion. Furthermore, some very rare CSCs are found to have both epithelial and mesenchymal state at the same time, representing the rarest CSCs with highest metastatic potential. The MET CSCs are identified by the expression of the aldehyde dehydrogenase (ALDH) isoforms, and the EMT CSCs are identified by the expression of CD44+ and CD24- expression.

To further validate and investigate the cancer heterogeneity, it's important to identify different cancer cell subpopulations to enable researchers to develop better treatments to eradicate cancer cells [19,20]. However, conventional assays usually only provide population-averaged results from a group of cells, which often masks important information from rare cell populations in cancer. To better probe cancer heterogeneity,

single-cell analysis is a promising strategy for researchers and clinicians to assay individual cells and study cancer heterogeneity at single-cell resolution [21,22].

1.2 Non-microfluidic single cell analysis methods and challenges

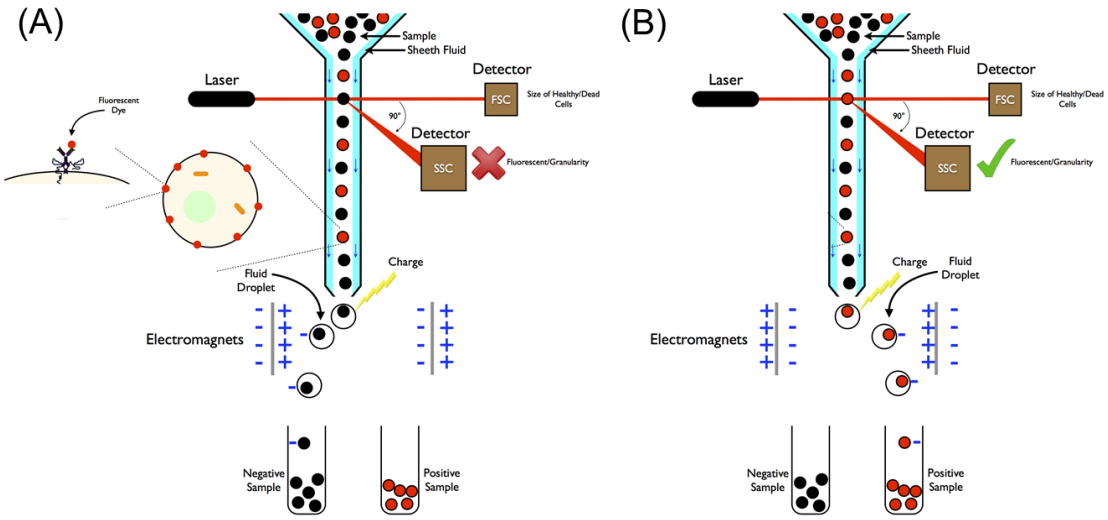

Figure 1-4 Flow cytometry (fluorescence activated cell sorting, or FACS) can analyze the biomarker with fluorescent tags at single cell resolution and sort single cells into different subpopulations for downstream applications. During the process, the fluorescent signal detector would determine if the passing cell contains the fluorescent tag and allocate the cells into two different tube accordingly for downstream analysis. The figure shows sorting cells with a negative signal (A) and a positive signal (B). [58] [59]

Single-cell analysis, however, is often a lot more challenging than conventional analysis due to the difficulty in single-cell handling and scale-down of assays to single-cell resolution. To fully probe cancer cellular heterogeneity, it's critical to achieve high-throughput to identify rare subpopulations that are critical for tumor progression such as the aforementioned cancer stem-like cells [8]. For single cell isolation, researchers have been using limited dilution to isolate single cells into well-plates by probability, but this method is labor intensive and also limited in throughput due to low single cell capture efficiency [23,24]. Single-cell picking is another isolation strategy to dispense cells into micro-wells and tubes for analysis. However, cell picking requires long imaging time to

identify the cell of interests and the serial picking procedure also limits its throughput. To overcome the throughput issue, some single-cell sorting methods such as flow cytometry can be used to facilitate high single cell capture rate. However, those methods are often expensive and the shear stress in the sorting process can significantly affect the viability and the assay outcome [25].

In addition, scaling assays down single-cell resolution presents another challenge for single cell analysis. For chemical sensing assays, limited analytes from a single cell diluted in microliter volume challenges the sensitivity of conventional assays that usually test analytes from thousands of cells [26]. Although there are some methods such as polymerase chain reactions that can amplify signals from low concentration for single cell analysis, high volume reagents consumption with the macro-scale assays quickly adds up the cost when scaling up the number of cells for analysis [27,28]. As a result, tremendous progress has been made to engineer microfluidics devices to simplify the cell manipulation process and enable chemical assays with miniaturized reaction volume.

1.3 Recent research of microfluidic single cell analysis

Microfluidics emerged as an attractive method for single cell analysis because of several technical advantages. First, by engineering flow channel structures, sensors and actuators at microscale, single-cell manipulation can be achieved as a first step for single-cell analysis. A variety of microfluidic single-cell handling capabilities have been developed to manipulate single cells for downstream assays [29–32]. Those cell handling methods are often categorized into active and passive sorting methods. Active sorting methods using dielectrophoresis (DEP) or optoelectronic tweezers (OET) often offer better selectivity and can lead to high cell capture rate, but they also require more sophisticated

systems for operation and are often more limited in throughput [33,34]. Passive methods such as hydrodynamic cell capture can be used to process single cells at scale, but many of them suffer from low cell capture efficiency because many cells are lost during the capture process. Different single cell capture mechanisms can be incorporated depending on the desired application given the different technical advantages. Second, handling cells in small volumes such as nano-liter or even pico-liter scale is another advantage of microfluidics [35,36]. By encapsulating the cells and analytes in small volume, the higher concentration of the analytes can be achieved and assayed with well-developed assay kits. For example, given the limited analytes from single cells, scaling the assay volume from micro-liter to nano-liter or pico-liter can increase the concentration by 3 to 6 orders of

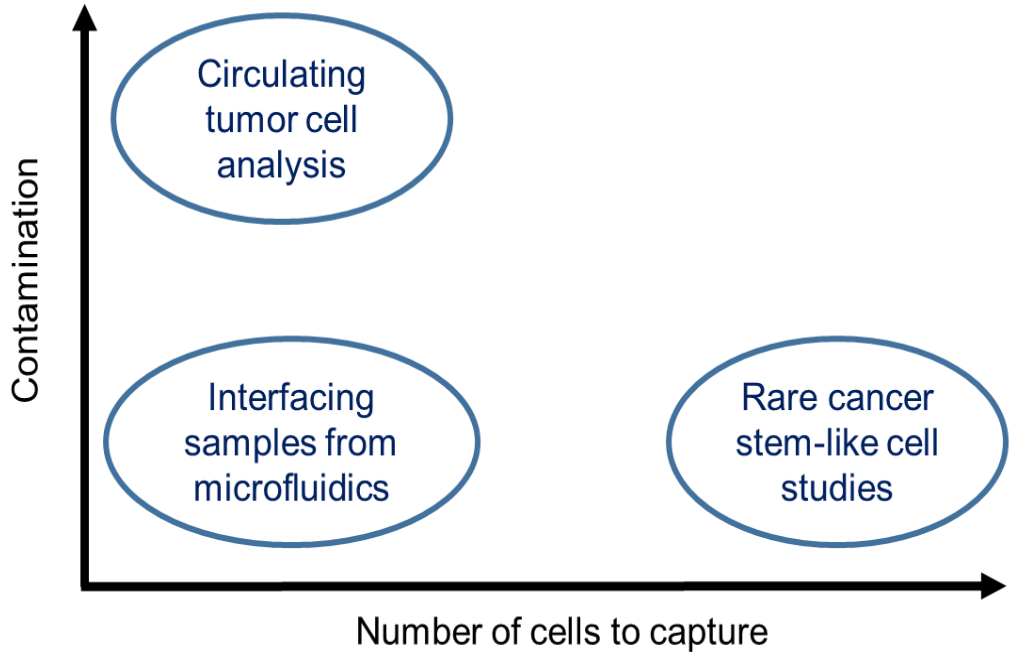

Figure 1-5 Highlights of the challenging samples with rare cancer cell populations addressed in this thesis. Chapter 2 presents a scaling and automation methods for single cell analysis to identify the rare CSCs from thousands of cells. As microfluidics become a popular method for cell property separation, Chapter 3 presents a method to interface samples from other microfluidics for downstream analysis. Chapter 4 presents the Hydro-Seq technology to achieve high-throughput and contamination-free single-cell RNA-sequencing of rare cell populations such as circulating tumor cells.

magnitude for better chemical detection. With small volume handling, microfluidics also reduces the consumption of reagent needed for the assay, which can increase the throughput of assays and reduce the reagent cost per cell. Last but not least, the integration of cell isolation and assay into the same chip, as a micro-Total-Analysis-System (MicroTAS), is another great advantage to enable seamless assay process flow from cell capture to on-chip analysis [37]. For instance, for single cell culture assays, researchers can directly engineer the microenvironment (co-culture environment or suspension environment) at the cell capture site. Therefore, right after cell capture, researchers can start observing cellular behaviors on-chip and validate different biological hypothesis without another cell manipulations [38]. For chemical sensing, chambers and valves can be designed to introduce reagents to the cell chambers to detect the analytes of interest [39]. After years of development in lab-on-a-chip systems, these platform technologies have demonstrated unprecedented capability to explore cellular heterogeneity.

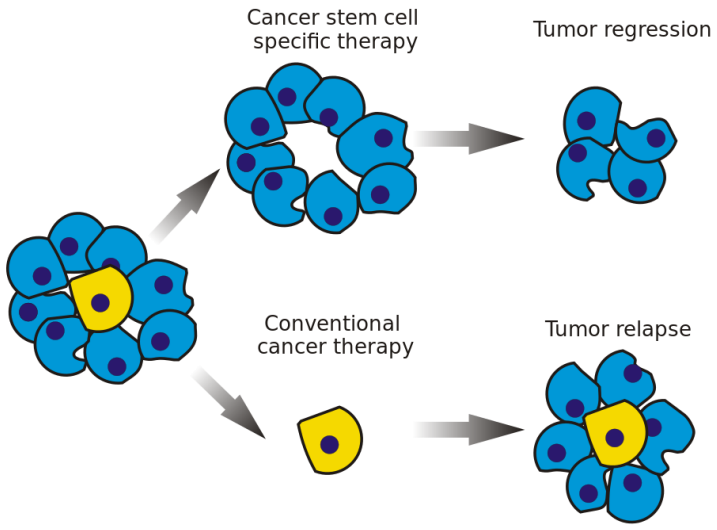

Figure 1-6 Cancer stem-like cells are a rare subpopulation of the tumor cells that leads to tumor relapse and tumor progression. Traditional cancer may shrink the tumor size, but the cancer stem-like cells can survive the treatment and regrow the tumor size after the treatment. As such, cancer stem-like cell specific therapies have been developed to treat tumors more effectively. [60]

1.4 Challenging samples with rare cell population for analysis

In cancer cell analysis, there are some important assays that only a small population of cells is available for study. Those rare cell populations can be only 1-5% of the bulk population as the cells of interests, so high-throughput single-cell capture is needed to attain enough cell population for reliable assays [38,40]. Another kind of rare cell population is the samples with only 10-100 cells of interests, so high cell capture efficiency to minimize cell lost will be a critical technical merit for successful experiments [9,41]. Although many microfluidic single-cell phenotypic and transcriptomic assays have been developed, it is still difficult to handle those challenging samples with rare cell populations [39,42]. In this thesis, I focus on interfacing three kind of rare cell populations for single cell analysis.

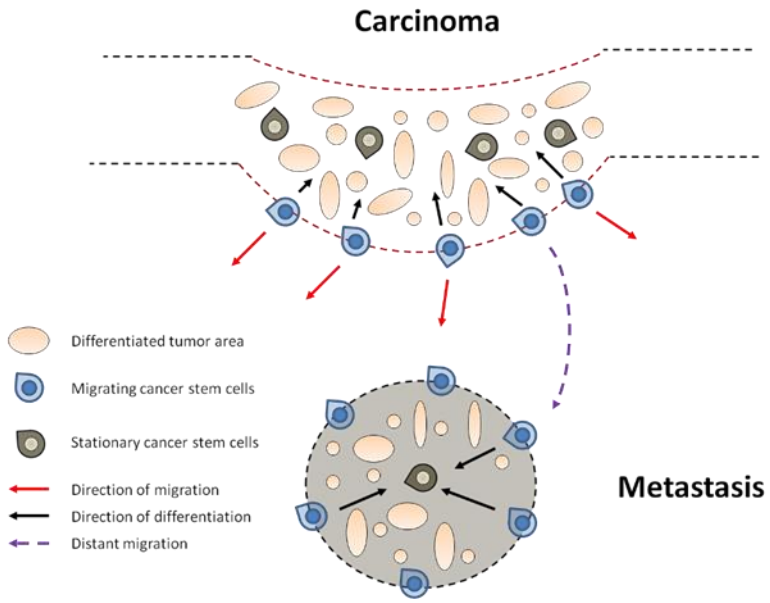

Figure 1-7 Tumor heterogeneity with cancer stem-like cells and the differentiated tumor cells. An important step for metastasis is the induction of epithelial mesenchymal transition (EMT) and mesenchymal–epithelial transition (MET) between the stationary and migrating cancer stem cells to enable cell invasion for forming a secondary tumor, which is the major cause of death for many cancer diseases. [61]

1.4.1 Studying rare cancer stem-like cells from bulk populations

Recent research suggests that there is a rare cell population driving tumor development, called tumor initiating cells (TICs) or cancer stem-like cells (CSCs), and this population only accounts for 1-5% of the total bulk tumor population[43,44]. Although CSCs can be studied and sorted by using some biomarkers tagged with fluorescent labels, it is possible that CSCs carry distinct marker expression from different tumors due to tumor heterogeneity and plasticity [13,17]. By utilizing the unique property of CSCs surviving under suspension culture and growing into single-cell-derived spheres, functional assays to culture single-cell-derived spheres is a promising alternative to study these rare cell populations [8]. As only 1-5% of the bulk population is CSCs, it is critical to capture sufficient cells (1000 cells or above) to start the sphere culture experiments, which will give rise to at least 10-50 CSCs and their spheres for analysis. Without sufficient cells, it will be difficult to quantify the sphere forming rate and thus obscure the CSC studies. To achieve reliable assays, there is an unmet need to develop a high-throughput cell capture and culture platform. Chapter 2 presents the scaling and automation of a high-throughput single-cell-derived sphere assay chip to address this problem.

1.4.2 Interfacing samples from microfluidic platforms

As microfluidic technology is becoming a popular platform for various cellular functional assays, there is a growing need to retrieve cells from microfluidics for further downstream analysis [41]. For instance, microfluidic single-cell migration chip has enabled functional sorting of motile and non-motile cells. By investigating the pathway regulation and molecular difference between these two populations, key regulation mechanism contributing to cancer metastasis can be identified for designing future cancer therapy [45,46]. In addition, single-cell-derived sphere assay also enabled the identification of CSC

population. After growing into tumor spheres, it is also important to characterize the intra- and inter-clonal heterogeneity of the CSC spheres and study the cell properties in the tumor microenvironment. Those samples from microfluidic could only provide a limited number of cells. Motile and non-motile cell population collected from microfluidic chips can range from only 10-1000 cells. A single-cell-derived sphere grown in microfluidic usually yields ~100 cells after dissociation. To interface those cells collected from microfluidic assays, it's critical to develop a cell loading scheme that can capture single cell effectively for analysis. Chapter 3 presents a proteolytic assay chip with a vacuum driven cell loading scheme to interface cell populations harvested from microfluidics.

1.4.3 Isolating circulating tumor cells for downstream analysis

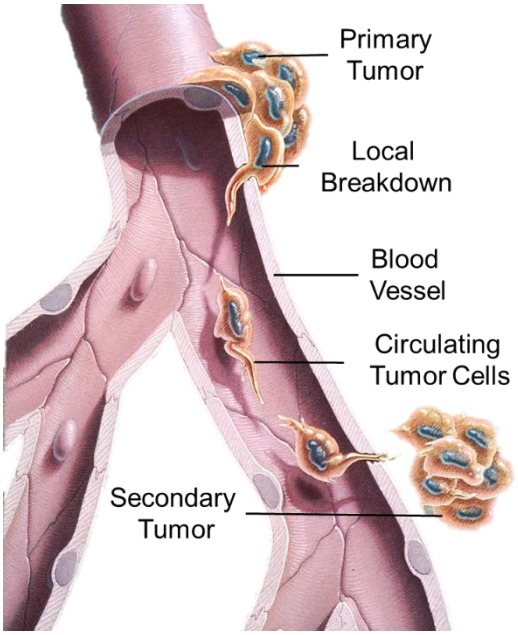

Figure 1-8 Diagram of cancer metastasis and circulating tumor cells. During tumor metastasis, the invasion tumor cells from the primary tumor migrates through the blood vessel into the circulatory system as "circulating tumor cells (CTCs)." To study the mechanism for metastasis, CTC molecular analysis presents a promising way to identify the metastasis-related pathway regulation and design new therapies to prevent cancer metastasis, which is the major cause of death for many types of cancer diseases. [62]

Circulating tumor cells (CTCs) are the tumor cells shed from the primary tumors into the blood vessels and metastasize into distant tissue locations in cancer patients [10,47,48]. With simple blood draws, liquid biopsy to collect CTCs provides a less invasive way to access to tumor cells as compared to regular biopsy by surgery. However, CTCs are extremely rare in blood. With 10mL of regular blood draw, only around 1-100 CTCs can be collected for analysis[47]. In addition to the rarity of CTCs, the excess amount of blood cells in the sample is another great challenge for CTC analysis. Given that there are billions of red blood cells and millions of white blood cells with a few CTCs in the sample, some described CTCs analysis like "finding the needle in the haystack." [49] To address the challenge, there were tremendous research to enable CTC enrichment [50–52]. With enrichment, most of the red blood cells and white blood cells can be removed to enhance the percentage of CTCs in the sample. To enable potential clinical diagnostic from CTCs, it is critical to further develop molecular assays for CTCs genetics, transcriptomics, and proteomics analysis. Some staining methods with fluorescent in situ hybridization (FISH) and immunohistochemistry (IHC) were developed, but they are limited to the number of fluorescent channels (~4 or 5) for multiplex analysis [52]. To isolation CTCs for multiplex molecular analysis such as RT-PCR and RNA-sequencing, single-cell picking methods including capillary suction and dielectrophoretic microfluidics have been used to isolate a few cells of interests labelled by fluorescent staining ell picking to individual tubes [53–55]. However, this method relied on fluorescent marker that can skew the sampling population, and the serial picking procedure is also very limited in throughput. Thus, there is an unmet need to develop a scalable high-single-cell-efficiency platform with the contamination washing capability. To address this need, Chapter 4 introduces the Hydro-

Seq technology to achieve high-throughput contamination-free CTC whole transcriptome sequencing.

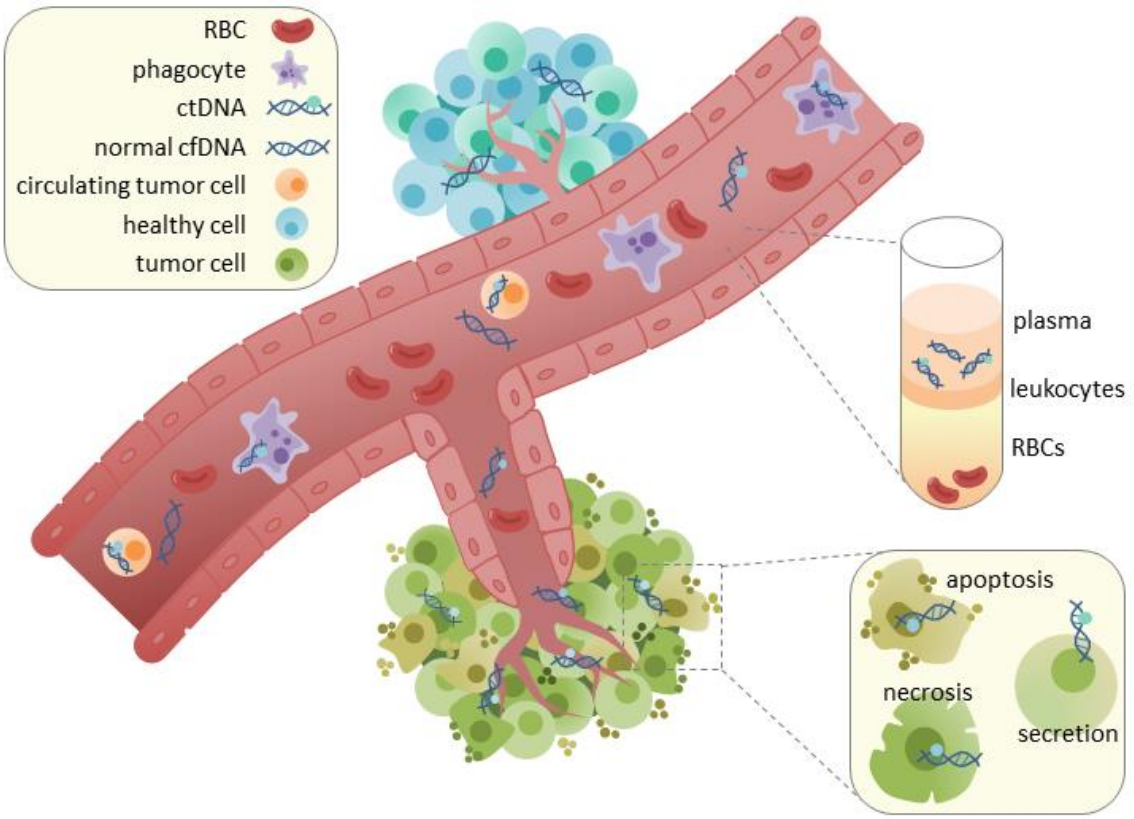

Figure 1-9 Cancer related biomarkers in the blood for liquid biopsy. As there are more cancer therapies targeting specific cancer biomarkers or subtypes, it's critical to identify the molecular profile of each patient for personalized therapy. By analyzing the circulating tumor cells and circulating tumor DNA (ctDNA) in blood, cancer molecular profiling with simple blood draws presents a promising way for precision medicine with the advantage of non-invasiveness compared to tumor biopsy [63]

1.5 Thesis outline

In this thesis, I focused on developing microfluidics to interface rare cell populations for phenotypic and transcriptomic assays. To enable small sample process capability, we focus on minimizing dead volume and achieving high capture efficiency without cell loss. To automate the assay readout, images processing program is developed to identify rare

cells for analysis. Since the assays are integrated on chip, the microfluidics were also designed to accommodate the protocols required in the assays.

Chapter 2 describes a high-throughput single-cell derived sphere formation assay chip with the focus of improvements on scaling and automation to study rare events from cell population. With a highly parallel chamber capture structure, the throughput of a chip can be scaled from 800 per chip up to 12,800 per chip with high single cell capture rate (~76.5%), which is over one order of magnitude throughput improvement from previous single-cell-derived sphere chips. In addition, a highly adaptive image analysis program was developed for cell/sphere counting and size measurement, enabling the automation of assay readout for high throughput analysis.

Chapter 3 introduces a single-cell proteolytic chip enabling the small sample loading with around ~50 cells. By a vacuum driven cell loading process, all the input solution can be loaded into the cell capture chambers to minimize cell loss in the dead volume. With the loading of protease sensitive reagent by diffusion and air isolation between each capture chambers, the proteolytic activity, an important process in metastasis, of each cell can be investigated with only few cells available.

Chapter 4 introduces the Hydro-Seq technology enabling single cell gene expression profiling for rare cell populations. Hydro-Seq captures single cells with high efficiency (>90% with only ~50 input cells) and pairs single cells with single barcoded beads for mRNA capture and sequencing, providing insights into the cell type composition. The utility of Hydor-Seq was demonstrated by sequencing CTCs from metastatic breast cancer patients, where 666 CTCs were whole transcriptome profiled with identification of critical biomarkers in metastasis and therapy.

Chapter 5 concludes the thesis with a summary of technical contributions and outlines the future research direction in both technical development and new application investigation.

Chapter 2
SCALING AND AUTOMATION OF HIGH-THROUGHPUT SINGLE-CELL-DERIVED TUMOR SPHERE ASSAY CHIP

Recent research suggests that cancer stem-like cells (CSCs) are the key subpopulation for tumor relapse and metastasis. Due to the cancer plasticity in surface antigen and enzymatic activity markers, functional tumorsphere assays are promising alternatives for CSC identification. To reliably quantify rare CSCs (1-5%), thousands of single-cell suspension cultures are needed. While microfluidics is a powerful tool in handling single cells, previous works provide limited throughput and lack automatic data analysis capability required for high-throughput studies. In this work, we present the scaling and automation of high-throughput single-cell-derived tumor sphere assay chips, facilitating the tracking of up to ~10,000 cells on a chip with ~76.5% capture rate. The presented cell capture scheme guarantees sampling a representative population from the bulk cells. To analyze thousands of single-cells with a variety of fluorescent intensities, a highly adaptable analysis program was developed for cell/sphere counting and size measurement. Using F108 coating on PDMS, a suspension culture environment was created to test a controversial hypothesis: whether larger or smaller cells are more stem-like defined by the capability to form single-cell-derived spheres. Different cell lines showed different correlations between sphere formation rate and initial cell size, suggesting heterogeneity in pathway regulation among breast cancer cell lines. More interestingly, by monitoring hundreds of spheres, we identified heterogeneity in sphere growth dynamics, indicating the

cellular heterogeneity even within CSCs. These preliminary results highlight the power of unprecedented high-throughput and automation in CSC studies.

2.1 Introduction

Cancer is known for its cellular heterogeneity, and it is believed that a small population of "cancer stem-like/initiating cells" (CSCs) is responsible for tumor metastasis and tumor relapse after treatment [1–5]. Some membrane surface markers and intracellular enzymatic markers have been used to identify CSC populations [6,7]. However, due to the heterogeneity and cellular plasticity of cancer, it is possible that CSCs carry distinct expressions from different tumors [8]. In addition to the markers, CSCs can also be recognized by their cell behavior [9]. When cultured in a suspension environment, CSCs can survive and proliferate into tumorspheres, while non-CSCs perform programmed cell death (anoikis) due to the loss of anchorage to substrates [10]. Hence, in-vitro single-cell-derived sphere formation assays are an attractive alternative to identify CSCs.

Performing in-vitro single-cell-derived sphere assays, however, is technically more challenging than traditional bulk assays. To ensure single-cell culture, researchers have used limiting dilution methods with low-attachment 96/384 well-plates to isolate single-cells in each well for sphere culture [11–13]. However, without a robotic system, this method is labor intensive and limited in throughput because the capture rate is limited by Poisson distribution (10-30%). Fluorescence-activated cell sorting (FACS) can automate the single-cell dispensing process and achieve higher single-cell seeding rate, but the high shear stress during the sorting can potentially affect cell viability and impact the results14. Given the low throughput of conventional approaches, people can barely quantify the sphere formation rate, so it is difficult to investigate the cellular heterogeneity within rare

CSC populations. The controversy of the correlation between cancer cell size and stemness is one example. In recent publications, some researchers reported evidence showing smaller cancer cell size is associated with cancer stem-like cell activity15–17, while other researchers reported higher mitochondria mass and increased cell size correlate with cancer stemness and chemo-resistance [18–20]. To study the heterogeneity in the CSC populations, there is an unmet need of an in-vitro high-throughput approach for rare cell studies.

Microfluidic culture systems emerge to be a powerful method for single-cell studies [21]. Combined with a non-adherent culture substrate, single-cell capture chips were developed for single-cell-derived sphere assay [22,23]. However, previous works with hydrodynamic capture system require extended meander channels to achieve a high cell capture rate (>80%), which constrained the number of wells per area [24–28]. The prolonged time for imaging over a large area limits the assay throughput and could potentially affect cell viability if an environmental chamber is not used with the microscope. For other capture methods, droplet systems can achieve high-throughput analysis by encapsulating single cells in aqueous droplets [29]. However, droplet approaches are limited by short assay time due to the difficulty in media exchange. Clonal sphere assay would typically require 14 days for culture and thus cannot be implemented by droplet systems. Micro-well systems are another simple yet effective tool to isolate single cells for clonal culture [30,31]. However, most micro-well systems rely on random seeding with a low cell capture rate around 10-30%. Schemes using dielectrophoresis force or dual-wells can facilitate higher capture rates, but they either require sophisticated active control or exhibit size-dependent capture [32,33]. High density and high capture rate can be realized by a filter array structure system [34,35]. However, in these works, capture sites were

connected in series, resulting in high flow resistance and low flow rate. Given that flow rate (2μL/hr), a high concentration of cells and a long loading time are required. In that case, cell aggregation and clogging may inevitably happen for many cancer cell lines. In addition, without automatic cell analysis, it requires manual inspection to read out data from microscope images, which is challenging when hundreds or thousands of single-cells are analyzed and the readout can be inconsistent when different people interpret the images.

To address these challenges, we have developed a high-throughput single-cell capture device utilizing highly-parallelized structures for single-cell-derived tumorsphere studies. The highly scalable fluidic structure enables reliable single cell capture from 800 wells/chip to up to 12,800 wells/chip. The capture scheme can reliably sample a representative cell population from bulk. With an automatic analysis program, assay results from thousands of cells and spheres can be analyzed after microscopic imaging. The high-throughput culture system with automatic analysis enables the analysis of heterogeneity within the CSC populations to study the cancer stemness – cell size correlation and single-cell-derived sphere growth dynamics.

2.2 Design of the high-throughput single-cell derived sphere

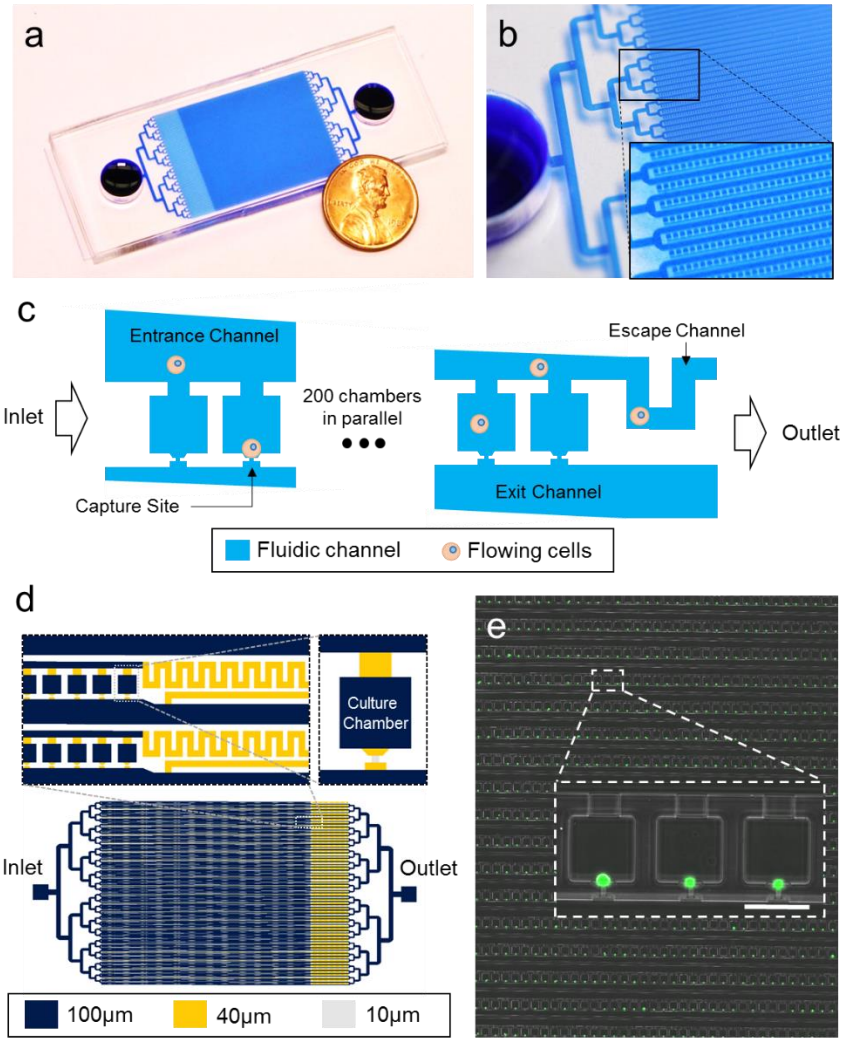

Figure 2-1 Overview of the high-throughput single cell culture chip (a) Photograph of a chip with 12,800 single cell chambers. (b) Close-up photograph of the branching channels and single cell chambers. (c) Schematic of a single branch channel with 200 single cell chambers. (d) Schematic of a whole device with specified structure height. (e) Microscopic picture of single-cells captured in the well arrays. (Scale bar: 100μm)

Single-cells are captured in micro-wells when they flow into the micro-wells and block the capture site. The micro-wells were designed to be 100×100×100μm3 cubes to provide room for sphere culture. To increase the throughput, it was found that simple duplication of identical micro-wells into a larger array will suffer from low cell capture rate and clogging caused by non-uniformity of cell distribution between the upstream and

downstream. To overcome this challenge, the scaling of the chip is achieved by engineering in two aspects. First, we investigated how to increase more cell capture wells in each branch channel. Second, the throughput can be scaled up by parallelizing the branch channels. A branch channel of 12,800-well chip is shown in Figure 2-1(c). In this design, each branch channel comprises an entrance channel and an exit channel with 200 micro-wells connected in parallel between them. After the entrance channel, a 40μm high escape channel was added to release residual cells in the entrance channel after loading. To ensure uniform cell capture at the upstream and downstream of each branch channel, we designed the unit flow resistance of entrance and exit channels to be significantly lower than each capture well by 100 times using multi-layer fabrication (Figure 2-2 and Figure 2-3). The flow resistance difference was achieved by implementing entrance channels and exit channels with a large channel height (100μm) compared to the capture sites (10μm). As the fluid was gradually transferred from entrance channels to exit channels when flowing to downstream, the entrance channels were tapered smaller and the exit channels were tapered larger to maintain the flow velocity in the channel. 40μm high channels were also used to connect the main channel to the micro-wells. When chips were flipped after cell capture, it formed a wall barrier around each micro-well to prevent cells from moving in and out of the wells (Figure 2-4). After finalizing the branch channel design, we connected the channels to the same inlet and outlet with branching channels. Using the scheme, chips with throughput from 800 wells/chip to 12,800 wells/chip are fabricated and tested. 800-well devices are composed of 16 branch channels with each containing 50 micro-wells, whereas 3,200-well devices are composed of 32 branch channels with each containing 100 micro-wells. (Figure 2-5) The 12,800-well array is composed of 64 parallel branch channels with each

containing 200 micro-wells as shown in Figure 2-1(a) and (b). This 12,800 single-cell well array covers a 24 mm × 27 mm area, so the whole chip with inlet, outlet, and other branching channels can fit on a 3'' ×1'' glass slide. The highly-parallel structure also results in low flow resistance, enabling gravity flow (100 Pa) cell loading by simple pipetting. The complete layout of a 12,800-well chip could be found in Figure 2-1(d). Figure 2-1(e) shows the microscopic image of cells being captured in the array, highlighting the power of high-throughput single-cell capture capability.

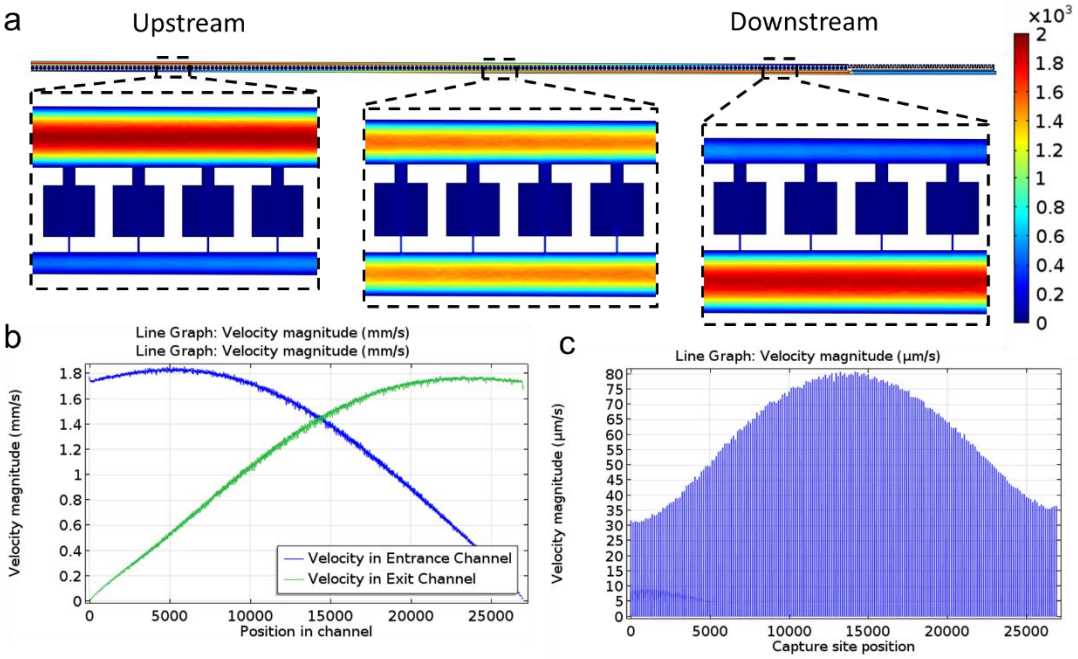

Figure 2-2 (a) Simulation of one branch channel with 50 Pa input pressure (unit: μm/min) (b) The tapered entrance and exit channel help maintain the constant flow rate at the upstream and midstream (c) Flow distribution across the 200 capture sites, showing higher flow rate at the midstream and lower flow rate at the upstream and downstream with ~65% difference from peak value. Although this distribution can be more uniform (<10% difference) by making the averaged cross-section of entrance channel larger, i.e. 150×200 μm instead of 85×100μm, this design is sufficient for robust single cell capture.

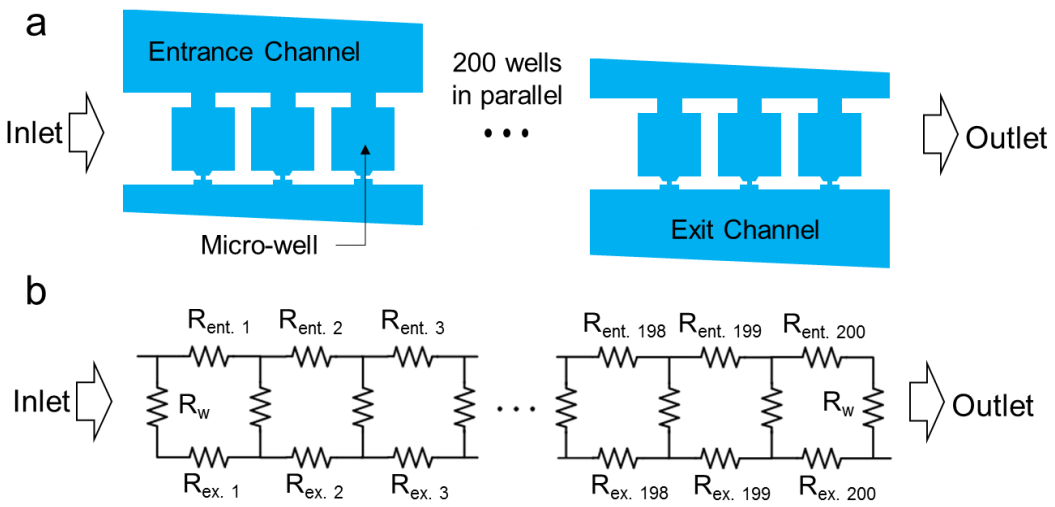

Figure 2-3 Fluidic circuit analogy for channel design. (a) Channel diagram of a single branch channel. (b) Equivalent electrical circuit diagram to the fluidic channel in (a). $R_{ent.}$ is the unit resistance of an entrance channel segment between two neighboring wells. The resistance gradually increases from $R_{ent.\ 1}$ to $R_{ent.\ 200}$ due to the tapered channel. R_w is the resistance through each micro-well. $R_{ex.}$ is the unit resistance of an exit channel segment between two neighboring wells. The resistance gradually decreases from $R_{ex.\ 1}$ to $R_{ex.\ 200}$ due to the tapered channel.

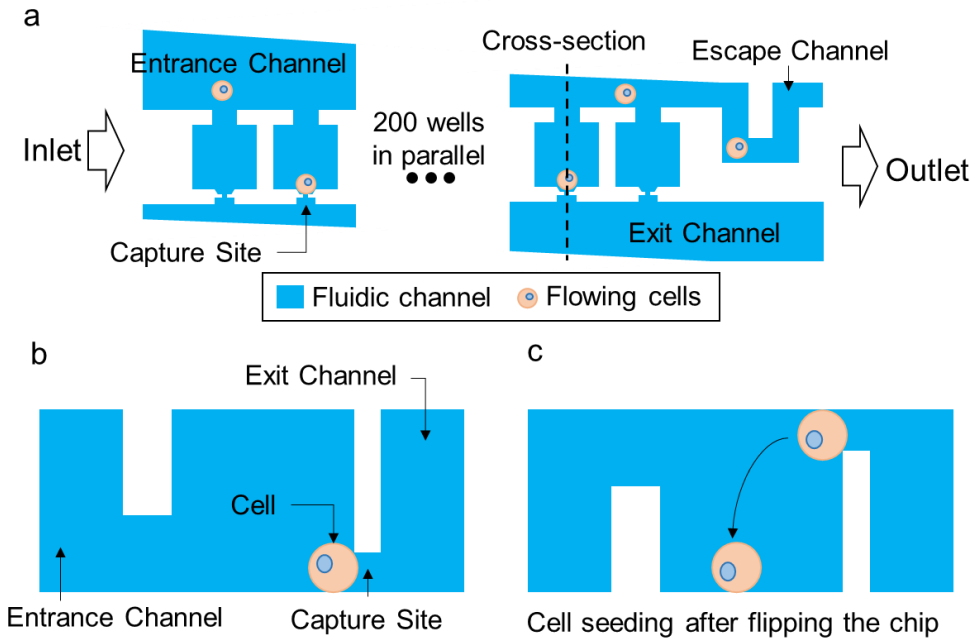

Figure 2-4. Cross-section diagram showing cell seeding after capture. (a) Overview of the fluidic channel with cells captured at capture sites. (b) A cross-section view of a captured cell from (a). (c) After flipping the chip, cells are seeded into each individual well by gravity for sphere culture.

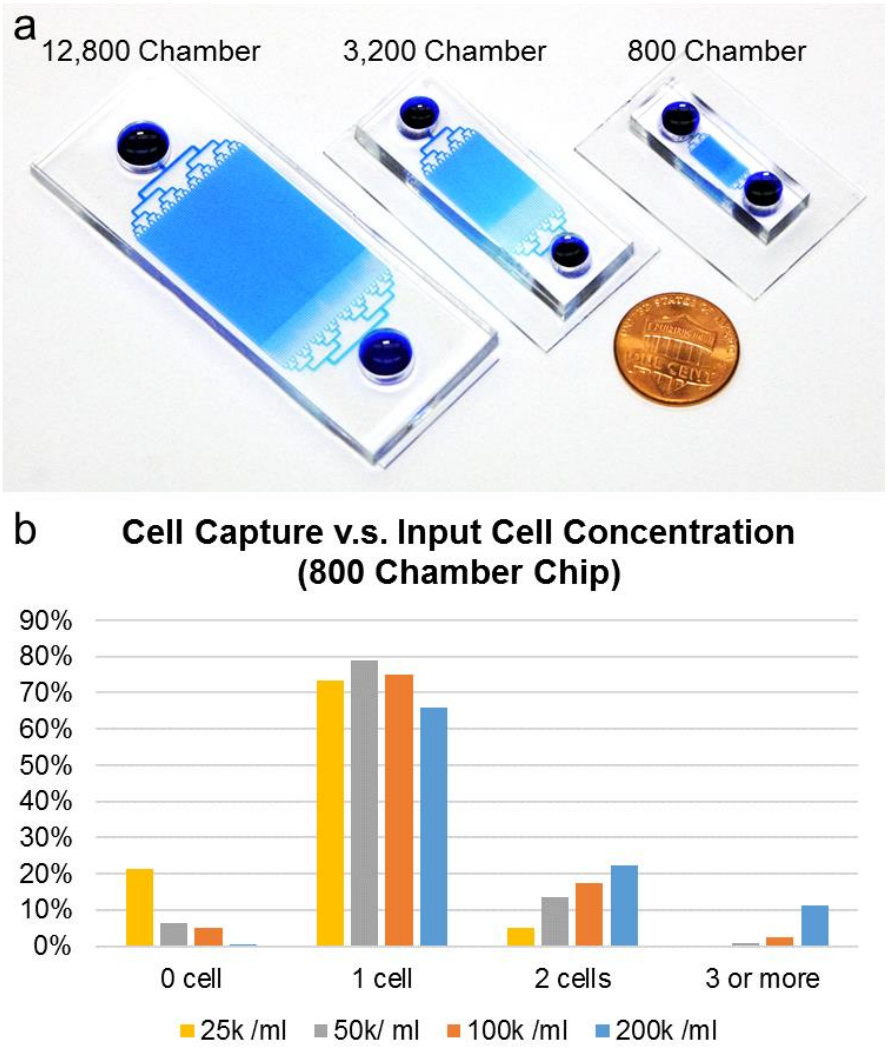

Figure 2-5 Device size of three different chamber numbers (a) Picture of devices with 800 chambers, 3,200 chambers, and 12,800 chambers for single cell capture and culture (b) Cell capture distribution for the 800 chamber chip, showing similar capture distribution compared to the 12,800 chamber chip. (N=1 for each concentration)

2.3 Automatic image analysis program

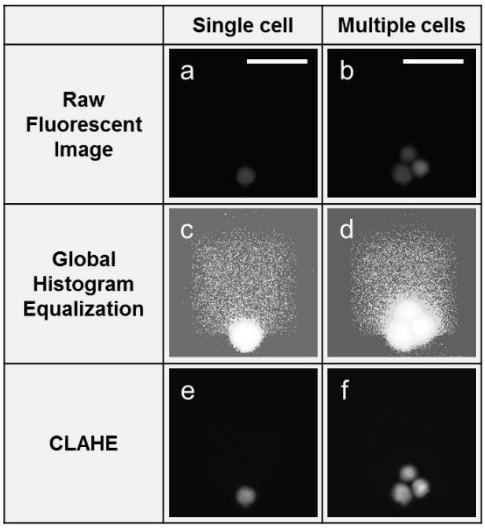

Figure 2-6 Contrast enhancement for image processing. Global histogram equalization creates artifacts from noise in the background, while CLAHE can reliably enhance the image of the fluorescent cells. (Scale bar: 50μm)

A custom-made MATLAB program, called μFAST, was developed to achieve "image-in-result-out" capability with manual sample checking function to ensure analysis quality. First, a user specifies the four corners of a chip, so μFAST can identify the location of each cell micro-well with a unique address using vector space operation. After well segmentation, μFAST performs image analysis to calculate the parameters of interest such as the number of cells and the size of the cell/sphere in each well. Due to the heterogeneity of cell fluorescent intensity, contrast enhancement should be done before analysis to ensure dim cells are also counted (Figure 2-6(a) and (b)). Since most parts of the background fluorescent image were dark, global histogram equalization generates artifact signals from noise in the background, making the image unusable for analysis (Figure 2-6(c) and (d)) [38]. To overcome this problem, contrast-limited adaptive histogram equalization (CLAHE) was used to enhance contrast in a localized patch area, minimizing the noise from a homogenous dark background (Figure 2-6(e) and (f)) [39]. For cell counting, a double k-

mean clustering algorithm was then applied to find the local peak of the fluorescent intensity to identify the position of cells (Figure 2-7 and Figure 2-8) [40,41]. A noise removal mechanism by morphological opening operations in local pixel area was also applied to remove false positives such as cell debris or local noise. This counting mechanism works reliability for both suspension and adherent cells (Figure 2-9 and Figure 2-10). To measure the cell size, the Hough transform was used to identify circular cells and their diameters (Figure 2-11) [42]. For sphere size calculation, intensity thresholding was applied to calculate the pixel area to extract the sphere area (Figure 2-12). The pixel area was then converted to μm^2 to calculate the sphere diameter. The presented µFAST program enables high-throughput analysis of thousands of cells with information such as cell size, number of cells, and also the sphere size for cell assay analysis.

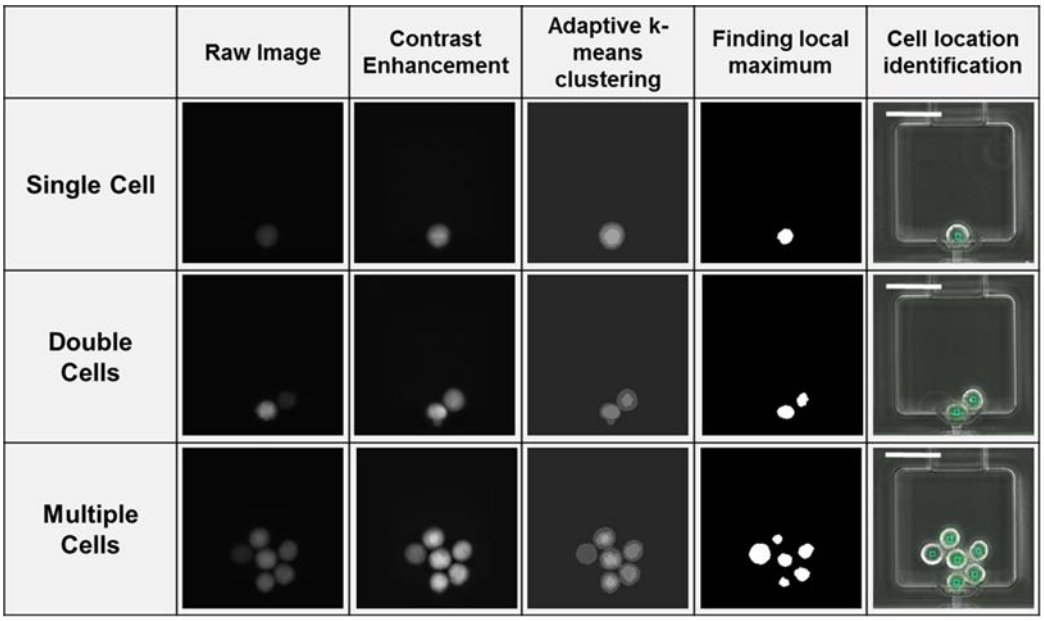

Figure 2-7 Cell counting with CLAHE contrast enhancement and adaptive k-means clustering to find the center of the cells in the image. Reliable counting results are achieved with examples of images with single cell, double cells, and multiple cells. (Scale bar: 40µm)

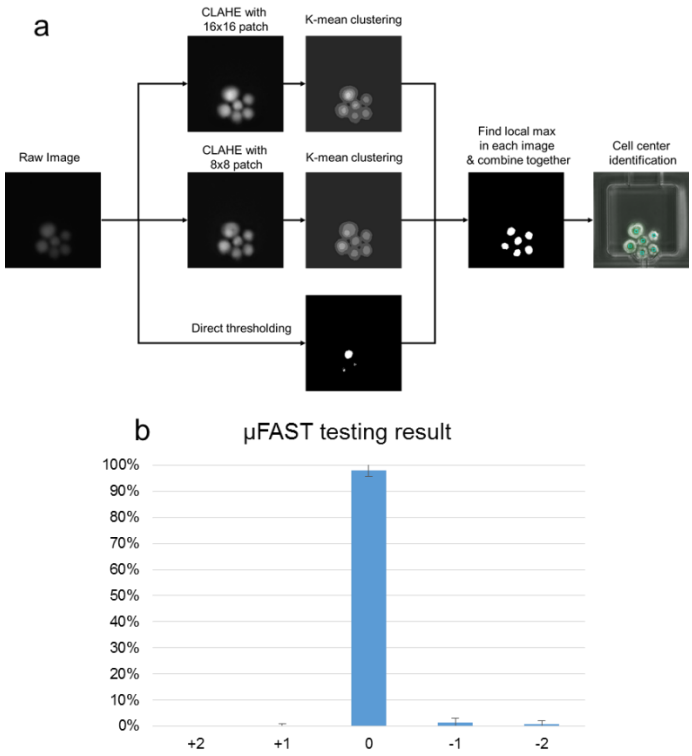

Figure 2-8 µFAST cell counting working flow and accuracy testing result (a) Step-by-step diagram of cell counting. (b) Example of accuracy testing result, showing 98% high accuracy with few cases of over-count (+1 and +2) and under-count (-1 and -2). (N=6 with total 350 micro-wells tested)

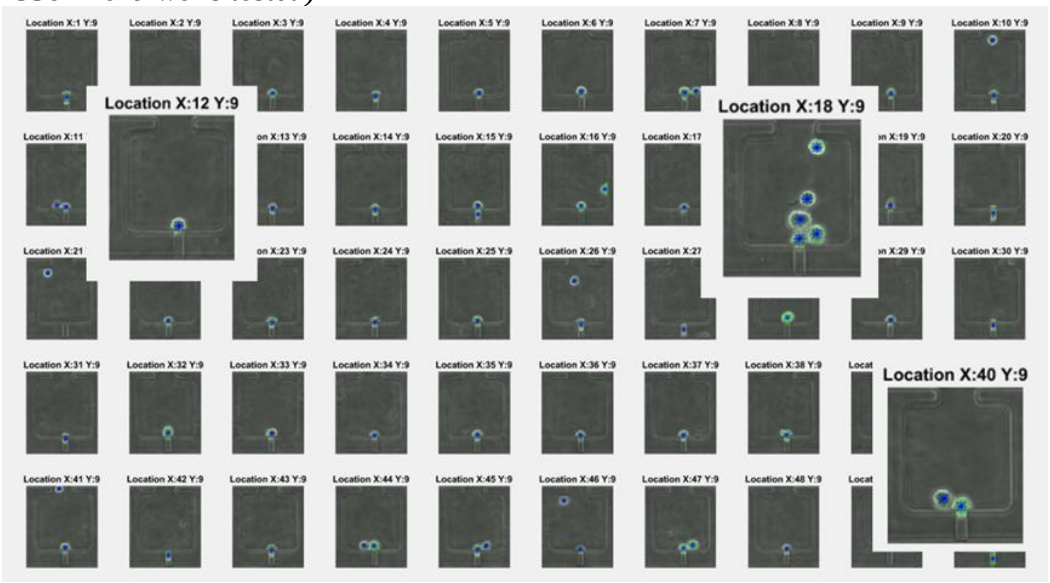

Figure 2-9 An example of cell counting panel after cell capture to confirm counting accuracy. Location X:12 Y:9 shows the single cell counting. Location X:40 Y:9 shows the case with two cells attaching to each other. Location X:18 Y:9 shows multiple cells counting.

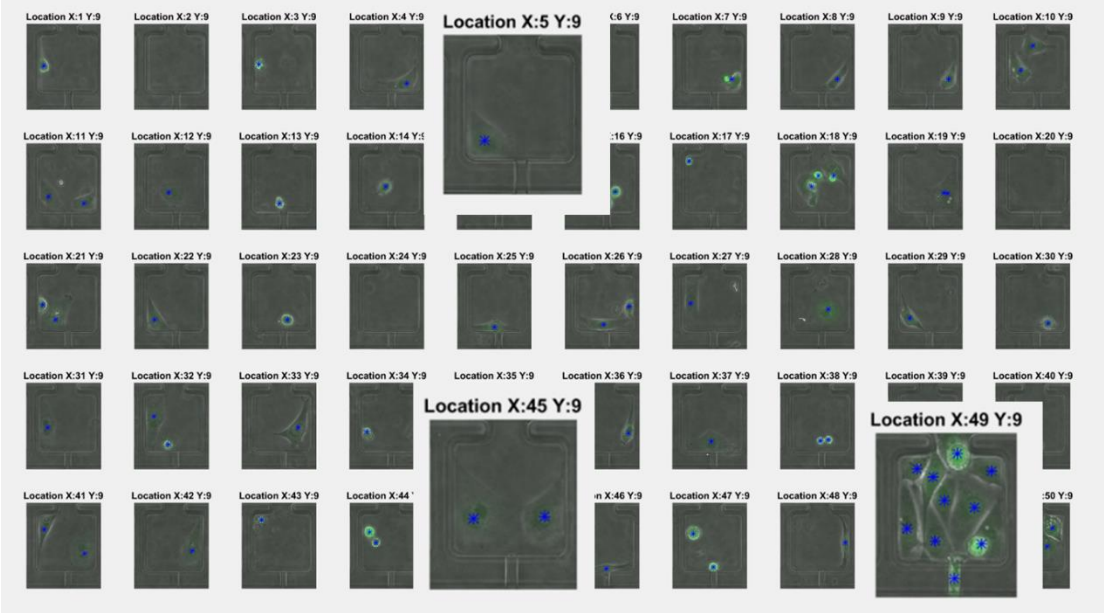

Figure 2-10 An example of cell counting panel for adherent culture. Location X:6 Y:9 shows single cell counting; Location X:6 Y:9 shows counting of two adherent cells; Location X:49 Y:9 shows cell counting of a confluent well, validating the counting accuracy of uFAST

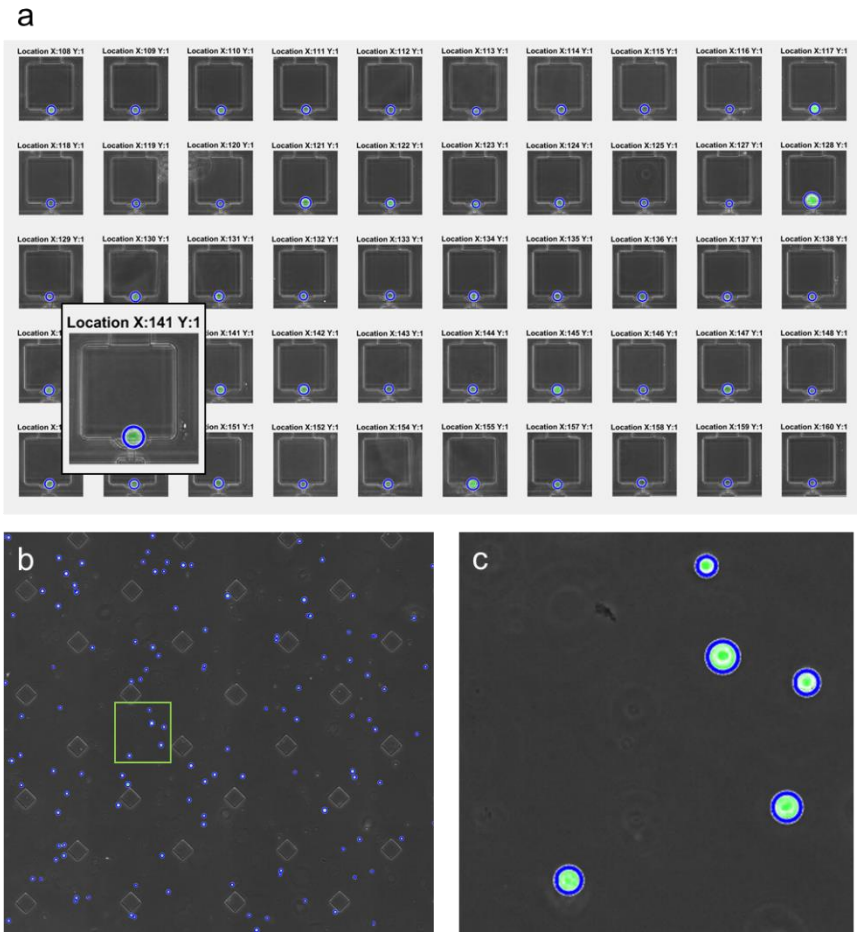

Figure 2-11 Cell size measurement with each cell size highlighted in blue circle. (a) An example of cell size measurement panel for cells captured on chip. (b) To measure bulk cell size distribution, cells were flooded into a 100µm-high chamber to ensure good imaging quality with cells on the same focal plane. (c) Cell size measurement of cells in bulk from the yellow box in (b).

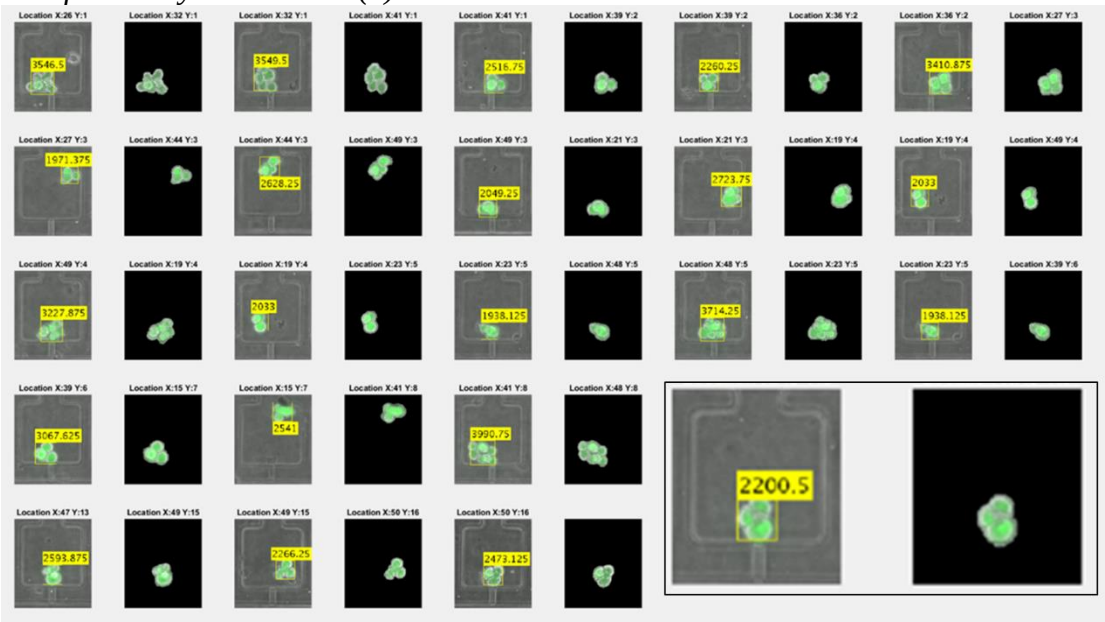

Figure 2-12 Sphere size measurement by fluorescent intensity thresholding

2.4 Cell capture and captured cell size characterization

We achieved a high single-cell capture rate of ~76.5% at the optimal cell concentration of 50k cells/mL. Reliable capture rate of >60% could be attained in a wide range of concentrations from 25-100k cells/mL (Figure 2-13(a)) Double and multiple captures have higher occurrence when higher concentration was used. It was observed that double and multiple captures have two different causes. They could be caused by the second cell coming into the micro-well before the first cell blocking the capture site. It could also be the consequence of the cells aggregating together in higher concentration in the cell suspension, resulting in cell cluster captures. Loading a lower concentration of cells helped reduce double and multiple capture significantly. However, the cell loading time is longer for lower cell concentrations, which can potentially affect cell viability. Given a 15 minutes

loading time, 50k cells/ml provides the optimal capture rate with consistent result. Due to the size heterogeneity of cancer cells, it's critical that the capture scheme could sample a representative population from bulk cells. The size distributions of cells captured on-chip and in bulk were measured and compared (Figure 2-13(b)). This shows the capability to capture cells with a wide range of sizes comparable to bulk cells as shown in Figure 2-13 (c-f). The high capture rate performance with a representative cell size distribution provides robust single-cell isolation to study highly heterogeneous cancer cells.

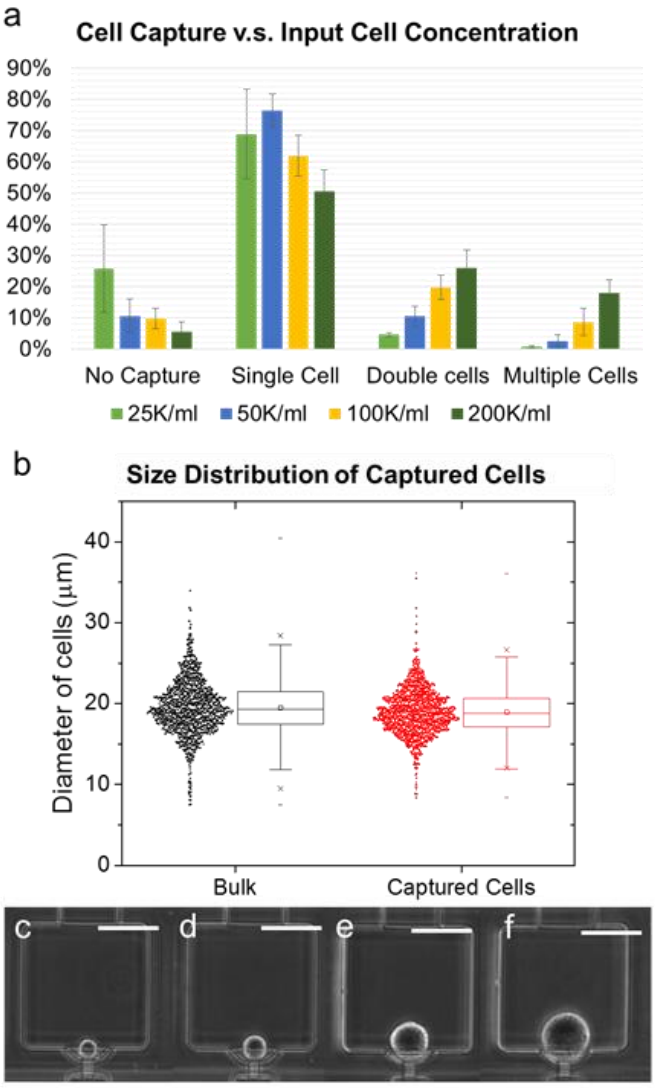

Figure 2-13 Cell Capture and Captured Cell Size Characterization (a) Cell capture result with different loading cell concentration (N=5) (b) Size distribution of MDA-MB-231 cells in bulk and captured in chip (N=600 for each case) (c)- (f) Examples of cells with different sizes captured in the chamber. The cell diameters are 12µm in (c), 20µm in (d), 30µm in (e), and 40µm in (f). (Scale bar: 50µm)

2.5 Single cell sphere formation rate and sphere size comparison

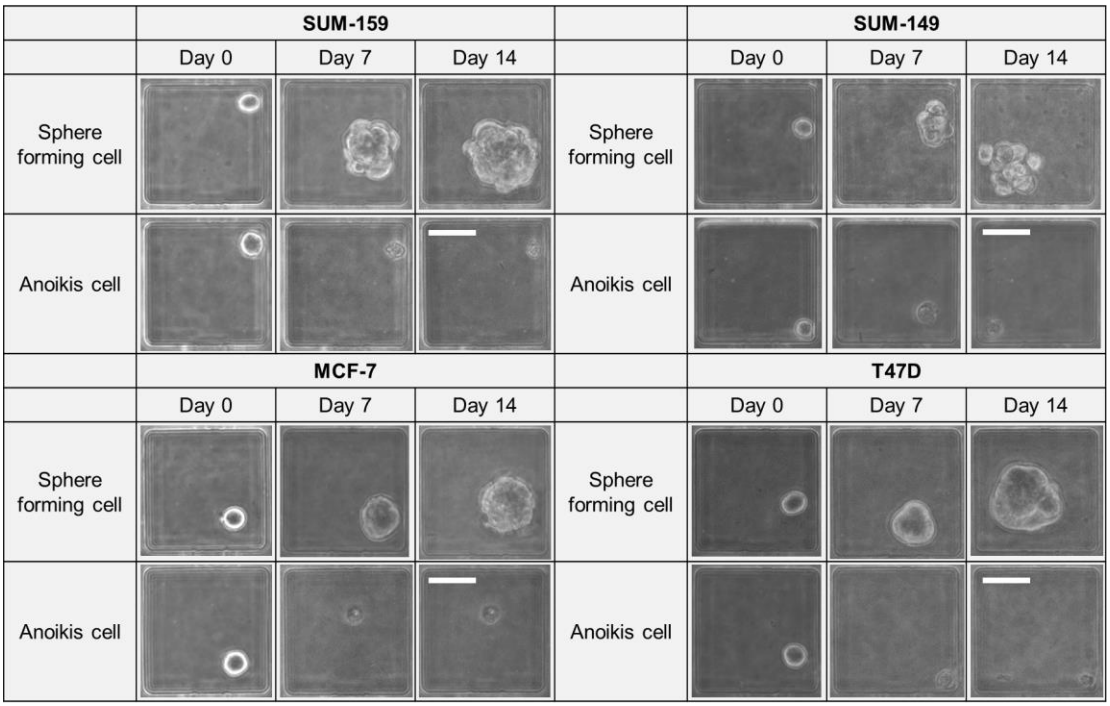

Figure 2-14 Microscopic images of cells in single-cell-derived assay with four breast cancer cell lines. Stem-like cells grew into a sphere from single-cells, while non-stem-like cells died of anoikis due to the loss of anchorage. (Scale bar: 40µm)

To validate the high-throughput single-cell-derived sphere assay, four different cell lines including SUM-159, SUM-149, MCF-7, and T47D were loaded and cultured in the device for 14 days. Since the PDMS was coated with F108, the polyethylene oxide (PEO) group prevents the cells from adhering to the substrate [36,37], enabling single-cell-derived sphere culture on chip to investigate sphere formation rates and sphere sizes. In suspension culture, stem-like cells grew into spheres from single-cells, while non-stem-like cells died as a result of anoikis (Figure 2-14 and Figure 2-15). After 14 days of culture, the sphere

size heterogeneity between cell lines and within cell lines was observed (Figure 2-15(a)). SUM-159 has the highest sphere formation rate around 45% and the average sphere size is also higher than the other 3 cell lines. Although SUM-149 has similar sphere formation rate to MCF-7 and T47D, the average sphere size is slightly higher than other two cell lines with statistical significance. The experiment reproduced similar sphere formation rates reported in previous *in-vitro* single-cell-derived sphere chips, validating the single-cell-derived sphere assay using this high-throughput platform [22,23]. As the sphere formation rate could be as small as 1~5%, this high-throughput system enables tens to hundreds of spheres to be analyzed from thousands of starting cells, allowing people to investigate the heterogeneity within the sphere populations.

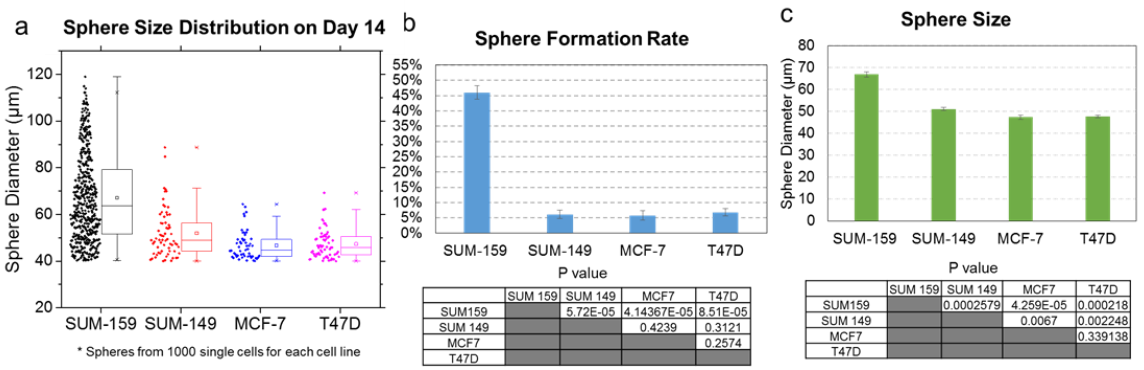

Figure 2-15 Large scale single cell derived sphere assay (N=3 for each cell line): (a) Sphere distribution after 14 days culture (1,000 single-cells randomly selected from each cell line). (b) Sphere formation rate after 14 days culture. (c) Sphere size after 14 days culture.

2.6 Sphere formation rate of different initial single cell sizes

After validating the high-throughput capability, we applied this technology to study the correlation of cancer stemness and cell size. Since the presented microfluidic chip is capable of capturing representative cell size populations, the correlation could be investigated by assessing sphere formation rates of subpopulations with different cell

sizes on the same chip. First, we measured the size distribution of the cells captured on-chip and separated them into two subpopulations based on cell size: small cells (bottom 30% in cell size) and large cells (top 30% in cell size) (Figure 2-16(a)). Then, the sphere formation rates of the two groups were compared in each cell line (Figure 2-16(b)). It was found that smaller cells have higher sphere forming potential than large cells in SUM-159, while SUM-149 and T47D have the opposite result. There was no significant difference in sphere formation rate observed between small and large MCF-7 cells. The result indicates that the relationship between cell size and stemness is cell line dependent, which may explain why different conclusions were attained by different groups.

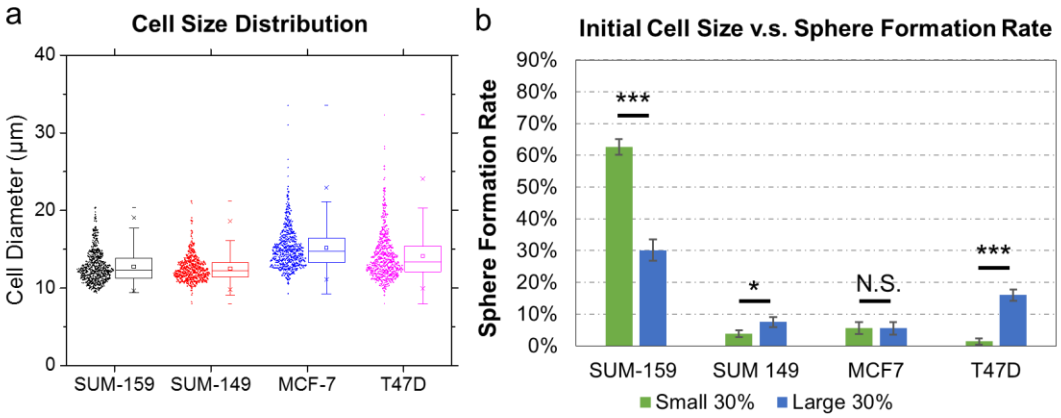

Figure 2-16 Comparison of sphere formation rate and the initial single cell size: (a) Cell size analyzed by μFAST showing different cell size distributions in cell lines. (600 cells randomly sampled for each cell line) (b) Sphere formation rate with small (bottom 30%) and large (top 30%) population from each cell line. (N=3 for each cell line)

2.7 Single-cell-derived sphere growth dynamics

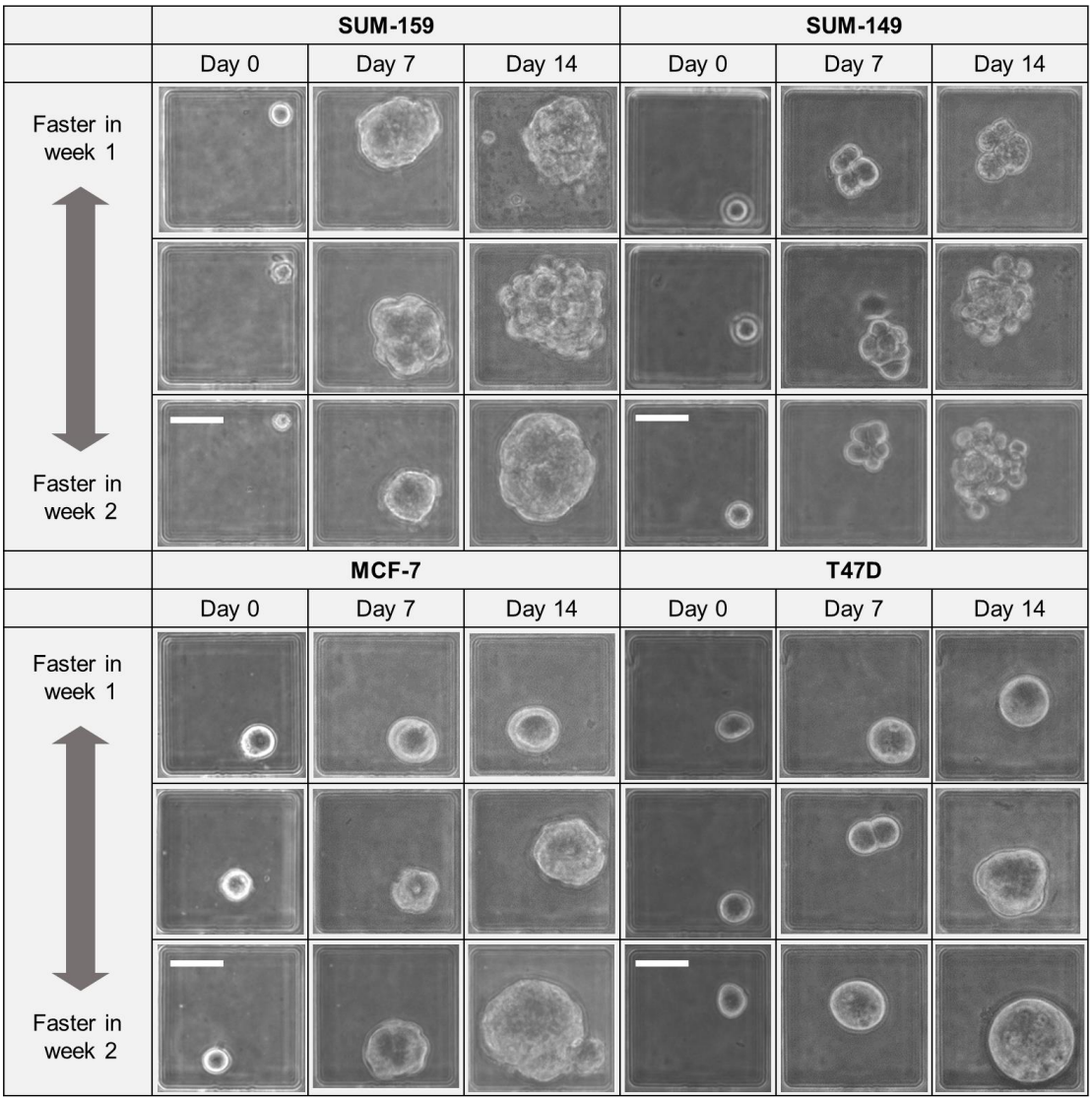

Figure 2-17 Microscopic images of cells showing different sphere growth dynamics over 14 days of cultures. In addition to spheres with steady proliferation, some spheres were more proliferative in the first week and became quiescent in the second, while others were more quiescent in the first week and became proliferative in the second week. (Scale bar: 40μm)

Another intriguing aspect of the presented sphere assay is the capability to track the sphere formation dynamics of each sphere with high-throughput, showing the proliferation patterns of the spheres (Figure 2-17, Figure 2-18, Figure 2-19, Figure 2-20, and Figure 2-21). For instance, in addition to the spheres growing at a constant rate, some spheres

could be more proliferative in the beginning and then became quiescent later, or vice versa. When comparing sphere diameter ratios from Day 14 to Day 7, heterogeneous sphere growth dynamic patterns could be observed. (Fig. 9 (a)) When the ratio is close to 1, it implies the limited proliferation of the sphere in the second week. When the ratio is larger, it implies faster proliferation in the second week. Among four cell lines, SUM-149 was observed to be more proliferative in the second week. To look into the difference of growth dynamics in SUM-159 and SUM-149, we plotted the sphere size on Day 7 and Day 14 on a scatter plot (Fig. 9 (b) and (c)). As shown in Fig. 9 (b), SUM-159 has wider spread due to the high proliferation rate and sphere size heterogeneity. In contrast, SUM-149 shows a cluster of samples within the 1.5× and 2× lines (Fig. 9 (c)). While all sphere forming cells are considered cancer stem-like based on conventional definition, huge variation of growth dynamics was observed, clearly implicating the cellular heterogeneity even within cancer stem-like cells. This observation can only be enabled by the scaling and automation of single-cell assay. In the future, this heterogeneity can be further investigated by using reporter to reflect the regulations within spheres and retrieving individual sphere for gene expression analysis.

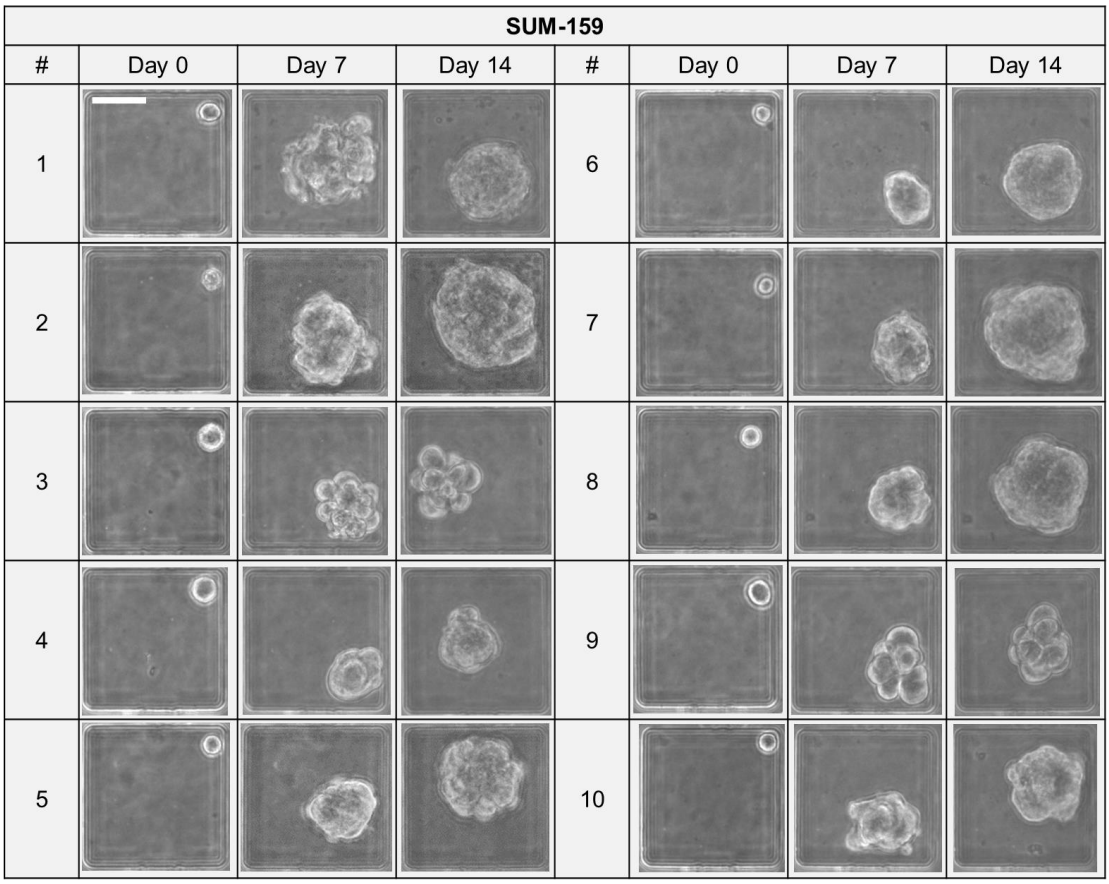

Figure 2-18 Microscopic images of SUM-159 single-cell-derived spheres (Scale bar: 40µm)

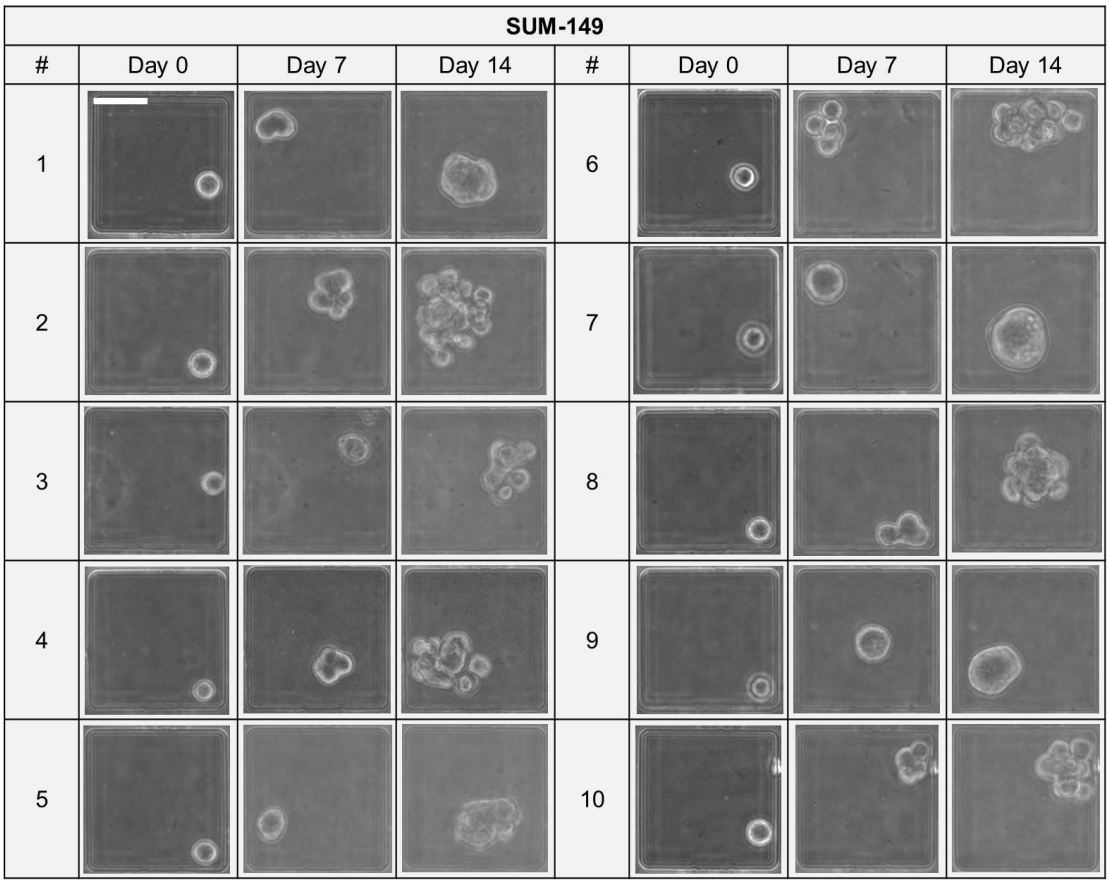

Figure 2-19 Microscopic images of SUM-149 single-cell-derived spheres. (Scale bar: 40μm)

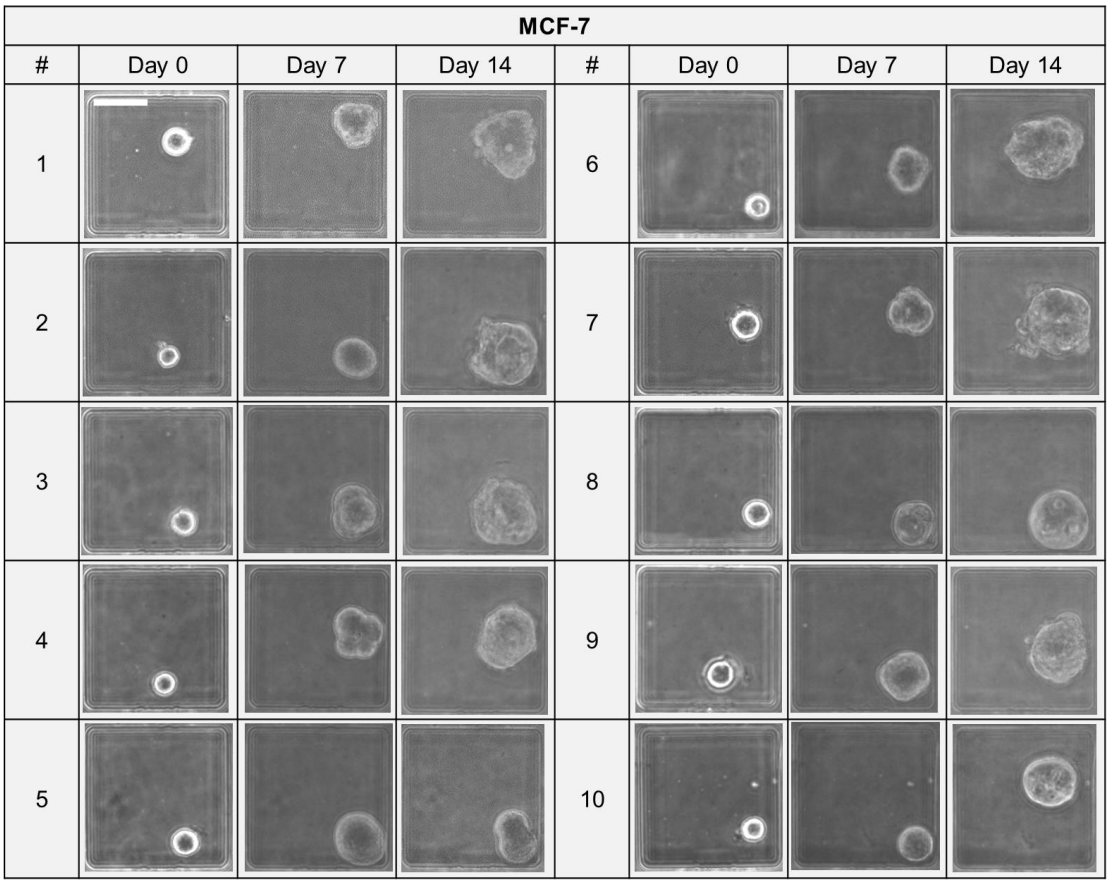

Figure 2-20 Microscopic images of MCF-7 single-cell-derived spheres. (Scale bar: 40μm)

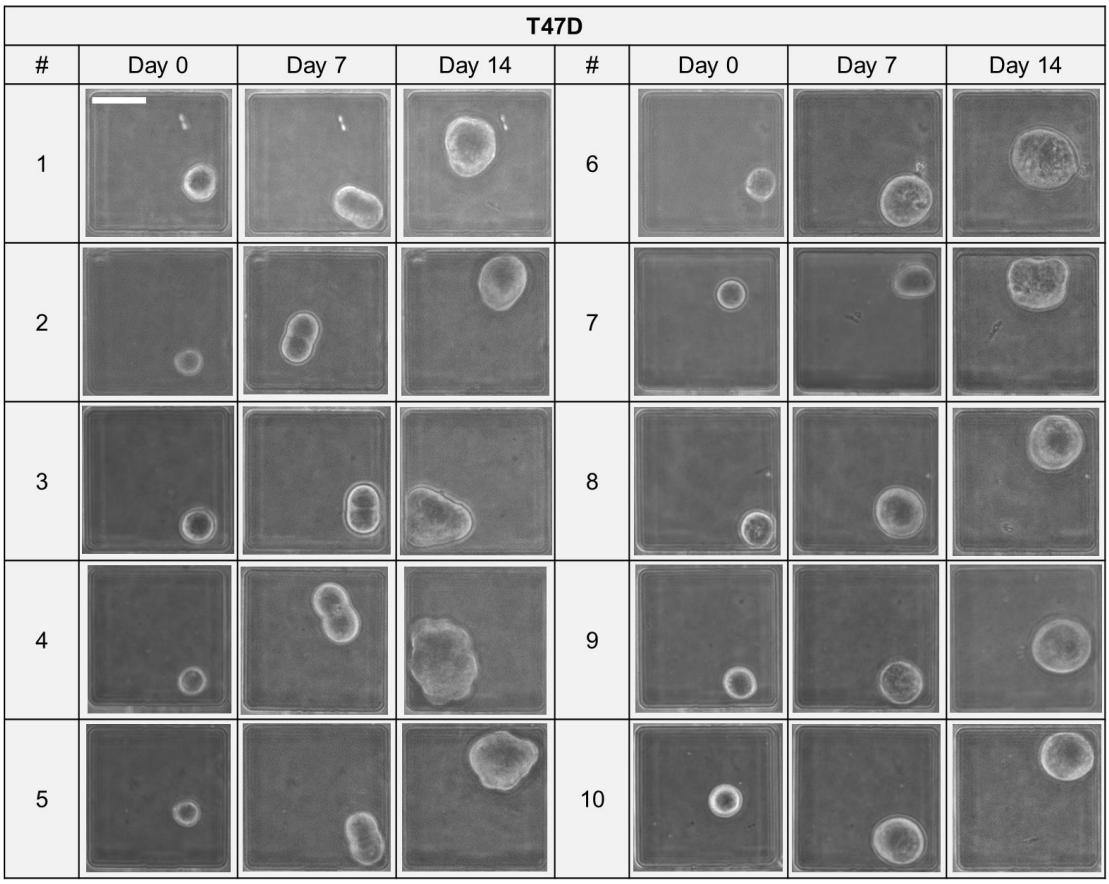

Figure 2-21 Microscopic images of T47D single-cell-derived spheres. (Scale bar: 40µm)

2.8 Chapter summary

We have presented a high-throughput analysis chip with highly-parallel structures for single-cell-derived sphere assays with wide range of scaling capability from 800 wells/chip to 12,800 wells/chip. For the highest throughput presented in this work (12,800-well chip), the optimal single cell capture rate (~76.5%) could be achieved with cell capture of a representative cell size population compare to bulk (10µm - 40µm in diameter). With the automated analysis software, µFAST, we were able to monitor various parameters of sphere assays in a high-throughput manner. First, we validated the single-cell-derived sphere culture and observed the sphere size heterogeneity across different cell lines. Second, we identified different cell size and cancer stemness correlations by evaluating sphere

formation rates of different size subpopulations. Finally, by monitoring hundreds of spheres, different sphere growth patterns were observed in different cell lines. The presented results demonstrate the power of a high-throughput approach with automatic analysis for the functional identification and analysis of cancer stem-like cell.

Chapter 3
SINGLE CELL PROTEOLYTIC ASSAYS TO INVESTIGATE CLONAL HETEROGENEITY AND CELL DYNAMICS USING AN EFFICIENT CELL LOADING SCHEME

Proteolytic degradation of the extracellular matrix (ECM) is critical in cancer invasion, and recent work suggests that heterogeneous cancer populations cooperate in this process. Despite the importance of cell heterogeneity, conventional proteolytic assays measure average activity, requiring thousands of cells and providing limited information about heterogeneity and dynamics. Here, we developed a microfluidic platform that provides high-efficiency cell loading and simple valveless isolation, so the proteolytic activity of a small sample (10-100 cells) can be easily characterized. Combined with a single cell derived (clonal) sphere formation platform, we have successfully demonstrated the importance of microenvironmental cues for proteolytic activity and also investigated the difference between clones. Furthermore, the platform allows monitoring single cells at multiple time points, unveiling different cancer cell line dynamics in proteolytic activity. The presented tool facilitates single cell proteolytic analysis using small samples, and our findings illuminate the heterogeneous and dynamic nature of proteolytic activity.

3.1 Introduction

90% of cancer-related deaths are caused by cancer metastases rather than the primary tumor [1]. Since proteolytic cleavage of extracellular matrix (ECM) proteins is essential in metastasis, the understanding of proteolytic activity can facilitate the design of new

protease targeting drugs for clinical use [2,3]. To investigate proteases such as matrix metalloproteinases (MMP) and adamalysins (ADAM), two important protease families in matric remodeling and growth factor shedding, researchers have developed protease sensitive fluorescent substrates based on fluorescence resonance energy transfer (FRET) mechanisms [4]. The fluorescence intensity of the dye increases when proteases cleave the amino acid-based substrate. As a result, the fluorescence intensity serves as a measure of proteolytic activity, enabling in vitro live-cell protease assays [5].

Due to genetic and epigenetic instability in cancer (caused by environmental factors, faulty repair mechanisms, etc.), subgroups of cancer cells in a tumor can have very distinct phenotypes, and these differences in behavior pose great challenges to the treatment of cancer [6,7]. Recently, researcher demonstrated that the cancer invasion is driven by the cooperation of heterogeneous cancer cells. A ‘‘division of labor’’ between inherently invasive cells, which possess protease activity, and non-invasive cells can facilitate tumor invasion. This research shows the importance of cell heterogeneity in proteolytic activity for metastasis [8,9]. As dish based methods only provide information about the average behavior of bulk cells, single cell resolution methodologies are required to unveil the mystery of tumor heterogeneity. In addition, cell dynamics is another intriguing aspect in oncology [10,11]. The study of cell dynamics can dissect the cell heterogeneity in the time domain, which can be critical for both fundamental cancer modeling and protease-related clinical solutions [12]. For instance, different treatment strategies can be implemented if only a small subpopulation of cancer cells have constitutively high proteolytic activity rather than all the cells going through cycles of high and low activity stochastically [6,13]. In order to probe cell dynamics, the capability to track an individual single cell

continuously is required [2]. As conventional dish-based methods do not provide methods for single cell tracking, single cell proteolytic activity dynamics has not previously been explored.

Thanks to their small sample handling capabilities, microfluidic technologies have already enabled single cell gene expression analysis, including real-time reverse transcription-polymerase chain reaction (RT-qPCR), digital PCR, and whole-transcriptome sequencing [14–16]. However, as proteases require enzyme activation to be functional, results may not reflect the true proteolytic cleavage activity [2,17]. Single cell western blotting allowing researchers to analyze proteins directly [18], but it is a destructive process, allowing only a single time point to be measured. Some microfluidic technologies that incorporated the aforementioned protease sensitive fluorescent substrates were reported for probing proteolytic activity directly [19], yet the existing tools have low cell loading efficiency while using small samples. However, since only a limited number of cells can be obtained from a variety of relevant samples such as CTCs, primary biopsies, microlavages, or when interfacing other microfluidic devices, the high efficiency in using small sample is necessary. In continuous-flow microfluidics, most single cell isolation processes, such as hydrodynamic or antibody-based capture, inevitably result in cell loss because of the dead volume and passive nature of the mechanisms, making it difficult characterize small samples [20–22]. For droplet based single cell assays, washing, supplying media, and assay substrate exchange is challenging, so it is infeasible to measure the proteolytic activity of the same cells at multiple time points to understand its dynamics [19]. In microwell-based systems, there are issues of media evaporation, reliable media

exchange, and microwell isolation [21]. Active capture mechanisms such as optical tweezers have low throughput and thus limit the utility of the technology [23].

In this chapter, we present a microfluidic proteolytic assay chip capable of capturing and isolating small cell samples and providing a robust methodology for media and reagent exchange. Using this platform, we investigated the heterogeneity that exists within cancer cell lines. Those that previously showed the importance of heterogeneity in invasion used a mix of cell lines to simulate innate heterogeneity; here we examined whether these characteristics are present within a single population and also their dynamic behavior. Additionally, by integrating two separate microfluidic approaches, we successfully examined inter- and intraclonal proteolytic heterogeneity. To the best of our knowledge, this is the first attempt to explore the clonal heterogeneity and dynamics of single tumor cells.

3.2 Single cell capture scheme

The presented platform is composed of a main-channel, which transports the cell suspension and sensing fluorescent substrate, an array of 1,000 chambers for single cell capture, and a vacuum channel, which is used to drive the solution into the chambers (Figure 3-1a). To measure the proteolytic activity, single cells and the commercially available FRET based substrate were loaded into the chambers. After loading, the chambers were isolated, and after the assay, the fluorescent intensity in the chamber indicates the activity of protease.

In order to facilitate high efficiency capture, a loading scheme has been developed that minimizes loading dead volume. In this scheme, the total volume of cell solution loaded is comparable to the total volume of the chambers (4 nL per chamber, ~4 μl total) and less

than the total chip volume. For loading, a solution containing the cells of interest (a couple uL in volume) is pipetted into the inlet and fills the first few rows of the main channel (Figure 3-1d,e). The solution does not fill the cell chambers due to surface tension effects and the air remaining in the chambers. Then, low pressure is applied to the vacuum channel that surround the cell chambers (Figure 3-2). As the material used for fabrication of this device (Figure 3-3), Polydimethylsiloxane (PDMS), has a high permeability [24,25], the air can diffuse through the sidewalls. In this way, the vacuum pulls the air from the chamber, and the air in the chamber can be gradually replaced by the cell solution within 90 seconds (Figure 3-1f-h). Once all the chambers within this first subsection are filled, the cell solution in the main channel can be pumped further downstream, and loaded into the next set of rows downstream (Figure 3-1i, Figure 3-4). After several iterations of these steps, the entirety of the cell solution sample can be loaded into the chambers throughout the device. In this manner, the scheme can minimize the dead volume to achieve efficient use of small sample, even down to around 10-100 cells. After loading, the distribution of cells per chambers should follow a Poisson distribution. When the number of cells loaded is much smaller than the total number of the chambers, it is likely that the chambers will capture single cells. As a demonstration, we loaded 10 cells in a 500-chamber device, and 7 single cells were captured in the chambers (Figure 3-5). When we loaded 100 cells, the distribution matched well with the Poisson model (Figure 3-1j). In cases where we have many cells, the number of cells loaded can be optimized using a Poisson model to achieve the highest capture rate (Figure 3-1k). After loading is complete, we can isolate chambers for the proteolytic assay by pumping air into the main channel. The isolation can be released by flowing media into the main channel for cell culture, providing simple and

robust valveless isolation (Figure 3-6 and Figure 3-7). Furthermore, the captured single cells were viable and proliferative even after 7 days of culture, indicating that the platform and the cell loading process do not affect cell viability and activity (Figure 3-8).

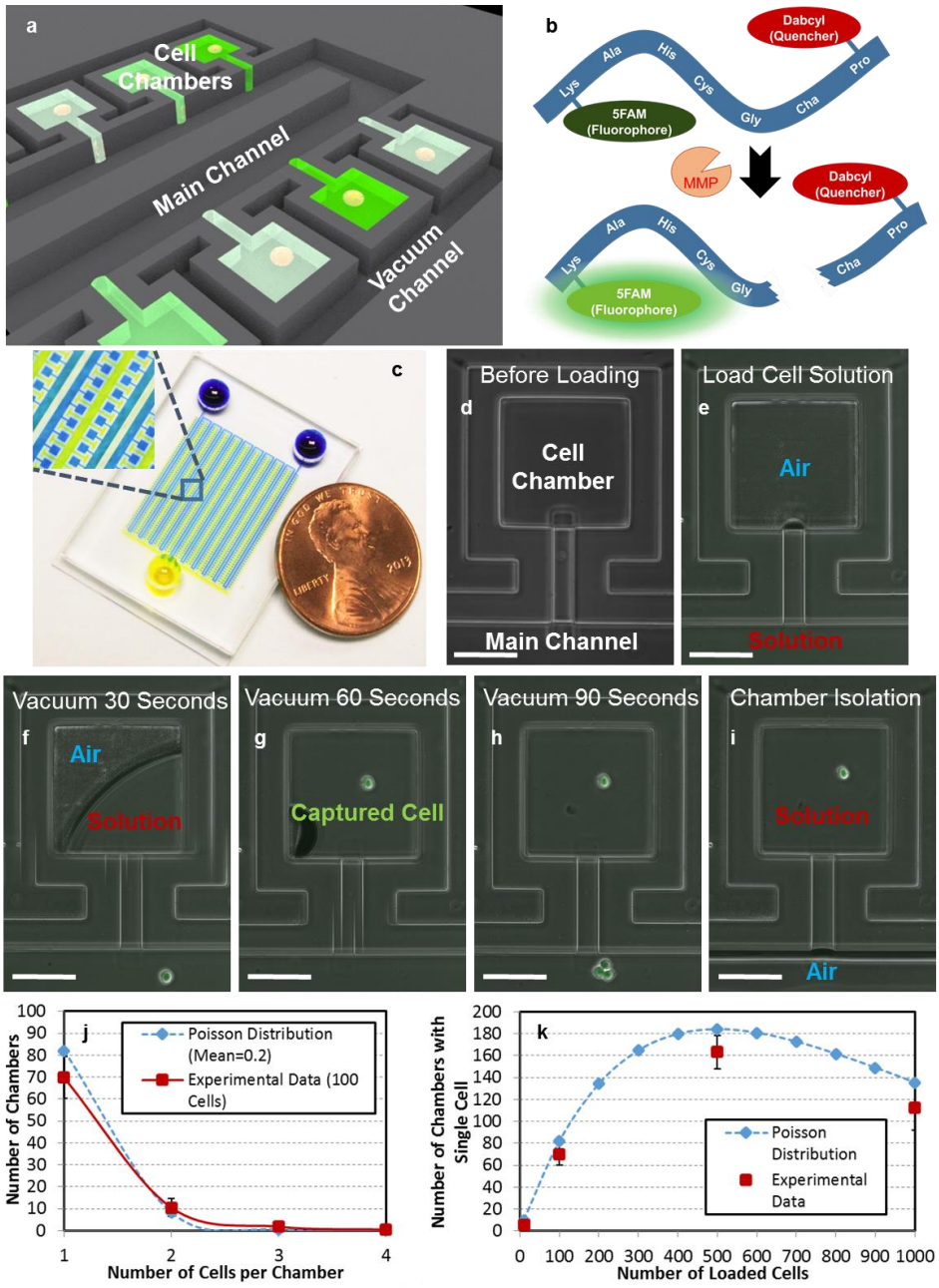

Figure 3-1 High-efficiency single cell capture scheme. (a) Schematic showing the single cell proteolytic assay. (b) The mechanism of the FRET-based sensing substrate. Before cleavage, the fluorophore is quenched. After cleavage, the substrate becomes brighter. (c) Photograph of a fabricated device. (d-i) The cell loading process: (d) empty chamber before loading, (e) loading the cell solution into the main channel, (f) solution partially fill

the chamber after 30 seconds of applied vacuum (g) after 60 seconds of applied vacuum, one cell was captured, (h) after 90 seconds of applied vacuum, the chamber was completely filled with the solution, and (i) the cell solution in the main channel was driven downstream for further loading. (scale bar: 100 µm) (j) The number of chambers capturing 1, 2, 3, or 4 cells, when loading 100 cells into a 500-well device (N = 4 devices). (k) The number of chambers capturing exactly one cell with different number of loaded cells (N = 4 devices). The results match well with the Poisson distribution model.

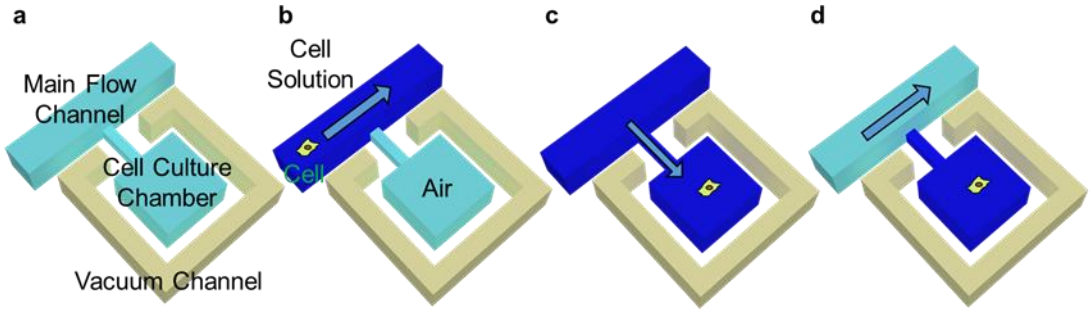

Figure 3-2 Vacuum driven single cell capture scheme. (a) The cell chamber is connected to the main channel and is surrounded by a vacuum channel. (b) In the beginning, the cell solution is driven into the main channel. (c) Vacuum is applied to the vacuum channel pulling the air through the PDMS and filling the chamber with the cell solution. (d) The cell solution is then driven further downstream to other chambers.

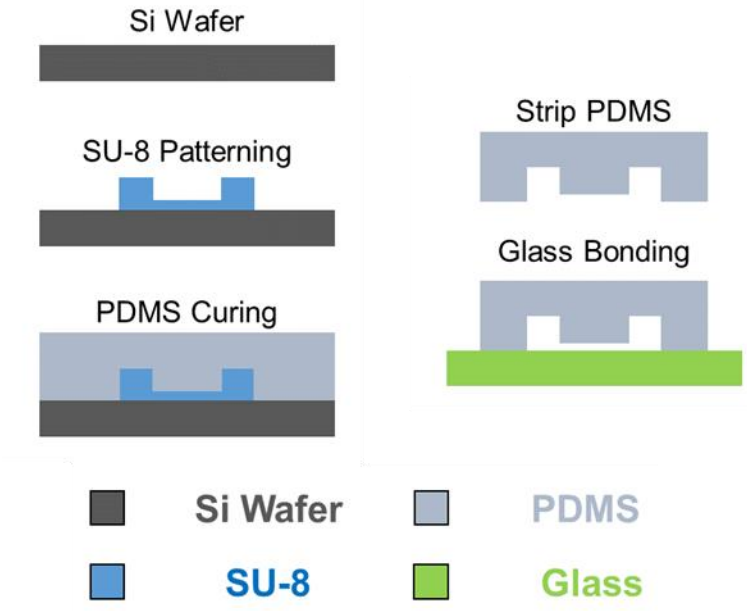

Figure 3-3 Device fabrication processes. The device is composed of a PDMS layer bonded on a glass slide. First, and SU-8 master was created using photolithography on a silicon wafer. Then, the PDMS layer was fabricated using standard soft lithography processes,

casting off the SU-8 master. Finally, the patterned PDMS was bonded to the glass after surface activation using oxygen plasma.

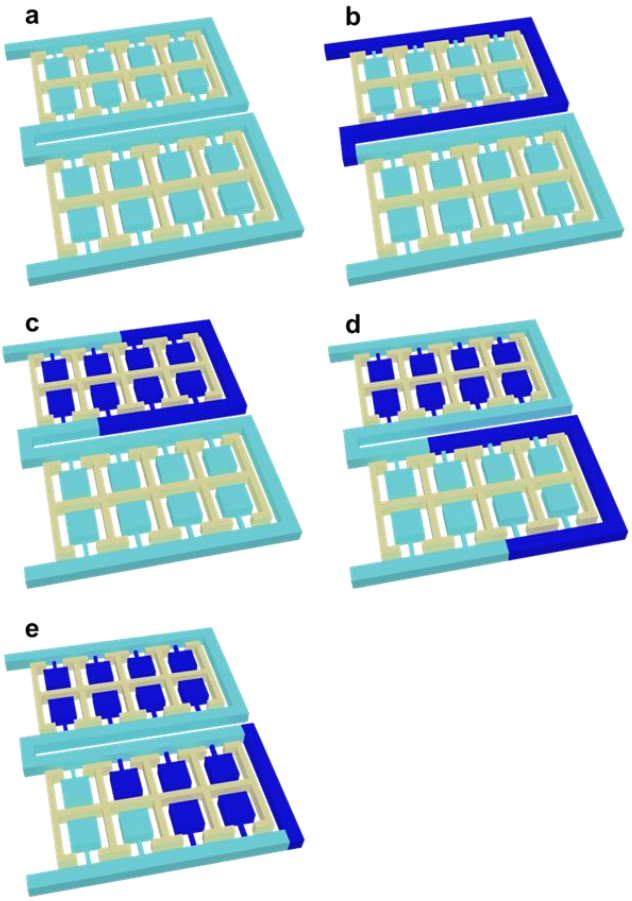

Figure 3-4 Serial high efficiency cell loading scheme. (a) The device consists of multiple rows of capture chambers that are loaded in a serial manner. (b) In the beginning, the cell solution is driven downstream into the first row. (c) Vacuum is applied to the vacuum channel, so the cell solution is driven into the chambers of the first row. (d) The cell solution is the driven further downstream to the second row of chambers. (e) Vacuum is applied to the vacuum channel, so the cell solution is driven into the chambers of the second row. After filling the chambers, the cell solution will be driven to the third row. In this manner, process is repeated until the total volume of the cell solution is loaded into the chambers of the device.

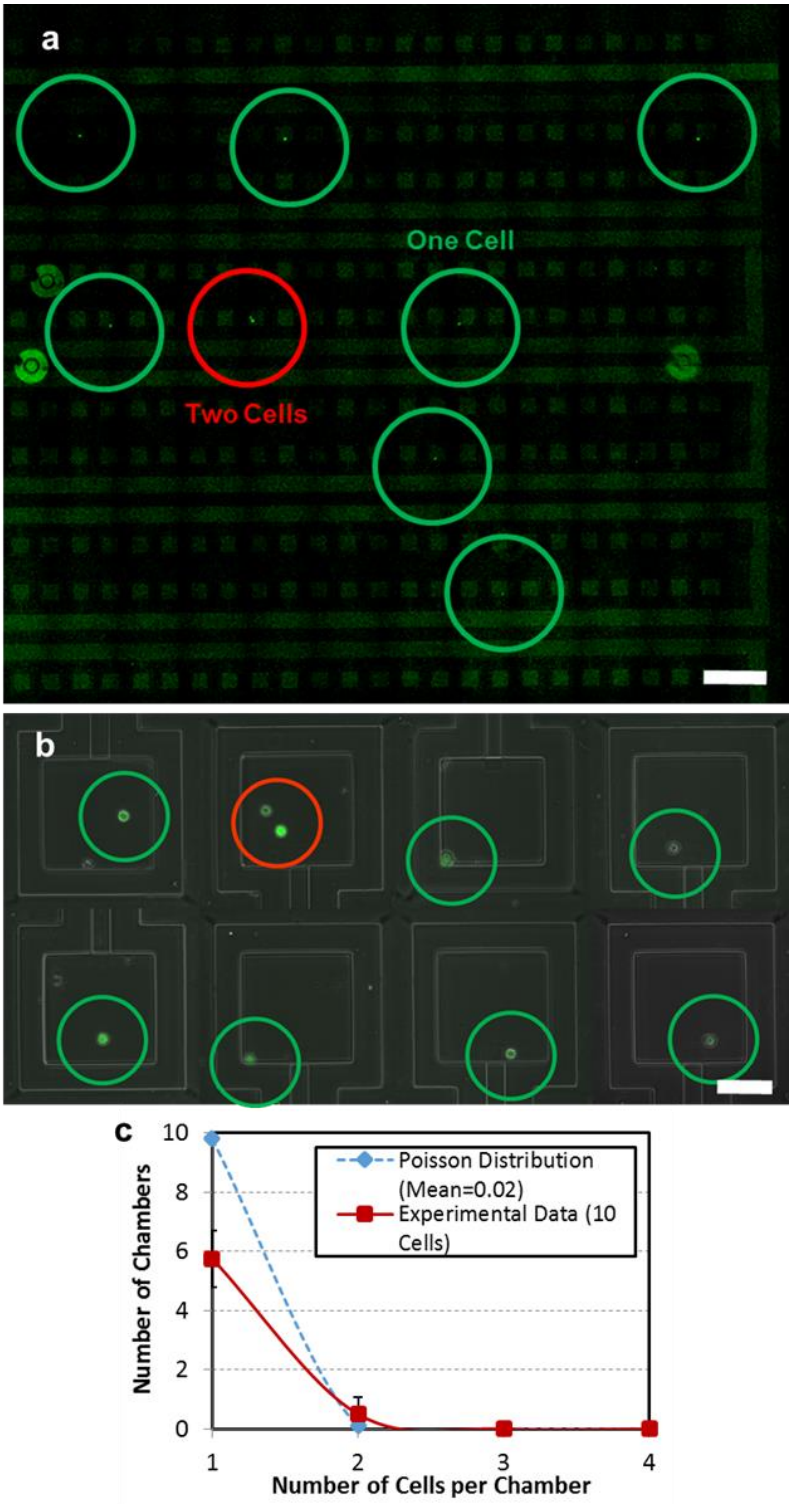

Figure 3-5 Cell capture of small samples (10 cells). (a) Fluorescent image of whole device (scale bar: 1 mm), (b) the enlarged view of chambers with captured cells (1-2 cells, scale bar: 100 µm), and (c) the Poisson distribution and the experiment results (N = 4 devices).

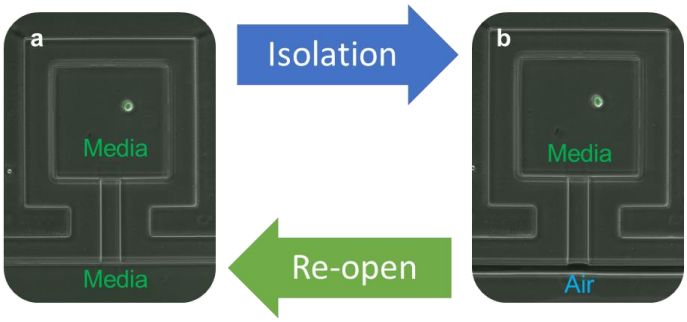

Figure 3-6 The isolation and re-opening of the cell chamber: (a) the chamber showing one cell captured before air isolation in the main channel and (b) after isolation.

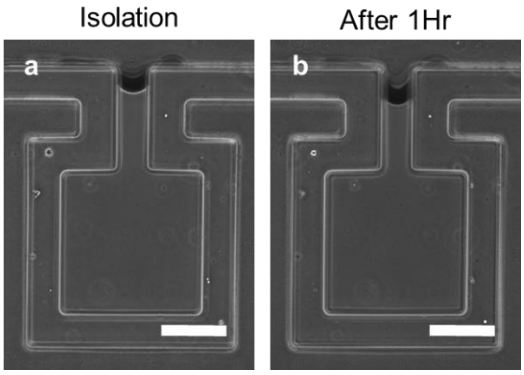

Figure 3-7 Evaporation of media during isolation. After one hour isolation, only 1-2% of the media has evaporated.

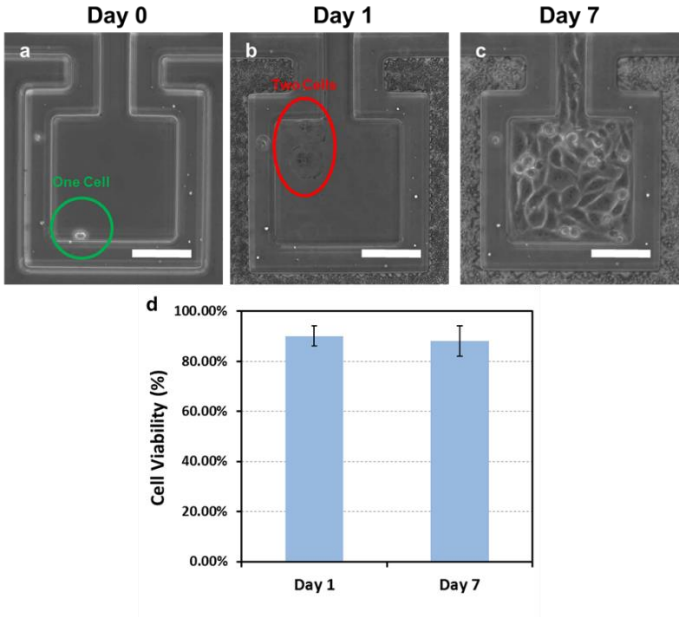

Figure 3-8 Cell viability of MDA-MB-231 in the chambers. (a) The cell chamber on Day 0 right after cell loading. (b) The cell chamber on Day 1. (c) The cell chamber on Day 7, showing high viability and proliferation on chip (scale bar: 100 µm). (d) The cell viability

was maintained at 90% for over 7 days, demonstrating that the platform has little effect on cell viability. (N = 4 devices)

3.3 Single cell proteolytic activity assay

To characterize the proteolytic activity of cells, a fluorescence resonance energy transfer (FRET) based substrate, composed of a FRET donor and quencher fluorophore, was used. The donor and the quencher are linked by amino acids, which can be cleaved by proteases (sequence: Dabcyl-Pro-Cha-Gly-Cys(Me)His-Ala-Lys(5FAM)-NH2). After cleavage, the distance between the pair increases, making quenching less efficient. Thus, the increase of fluorescent intensity can be used to indicate the proteolytic activity in the chamber.[56] First, we performed control experiments, verifying that fluorescent intensity increases with higher concentration of trypsin, which cleaves proteins, and longer assay time (Figure 3-9 and Figure 3-10). Also, the proteolytic activity of multiple breast cancer cell lines was measured in 96-well plates (Figure 3-11). These produced the expected results, validating the substrate and approach.

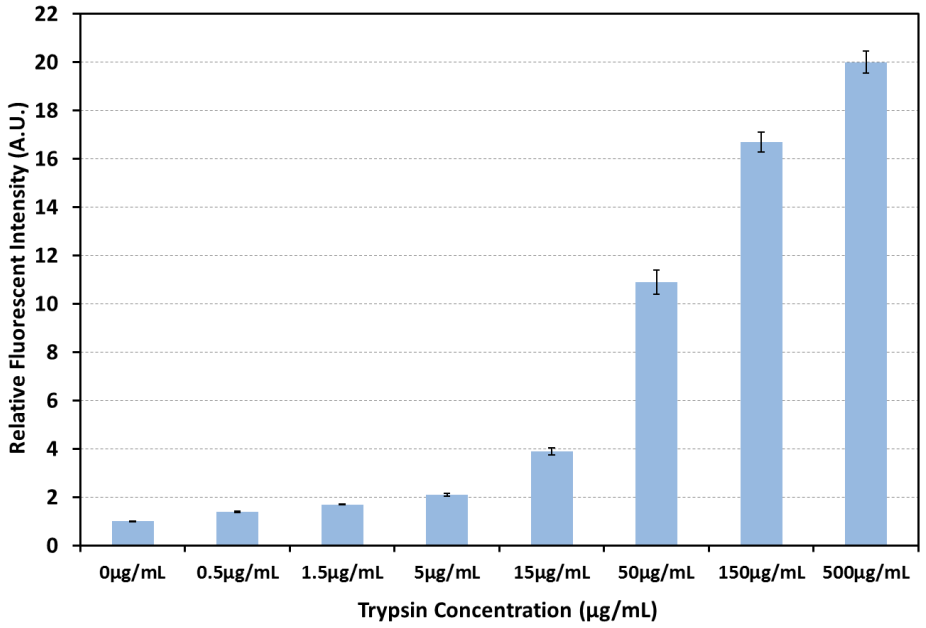

Figure 3-9 The fluorescent intensity of the substrate versus different concentration of trypsin (30 minutes reaction time), a protease that cleaves the FRET substrate. High concentration of trypsin cleaves more substrate, and the FRET signal becomes brighter. (N = 6 wells in 96-well plate)

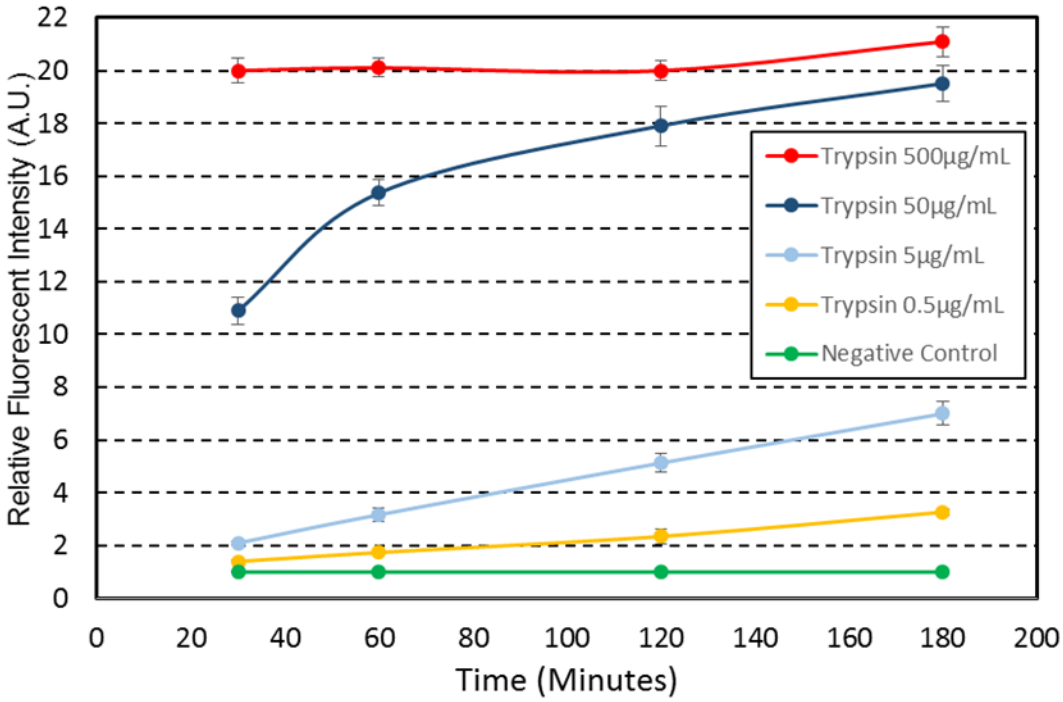

Figure 3-10 The fluorescent intensity of the substrate versus different reaction time. Under the conditions of low trypsin concentration, the fluorescent intensity increases with increasing reaction time. Under the conditions of high trypsin concentration, the intensity saturates within the first 30 minutes. (N = 6 wells in 96-well plate)

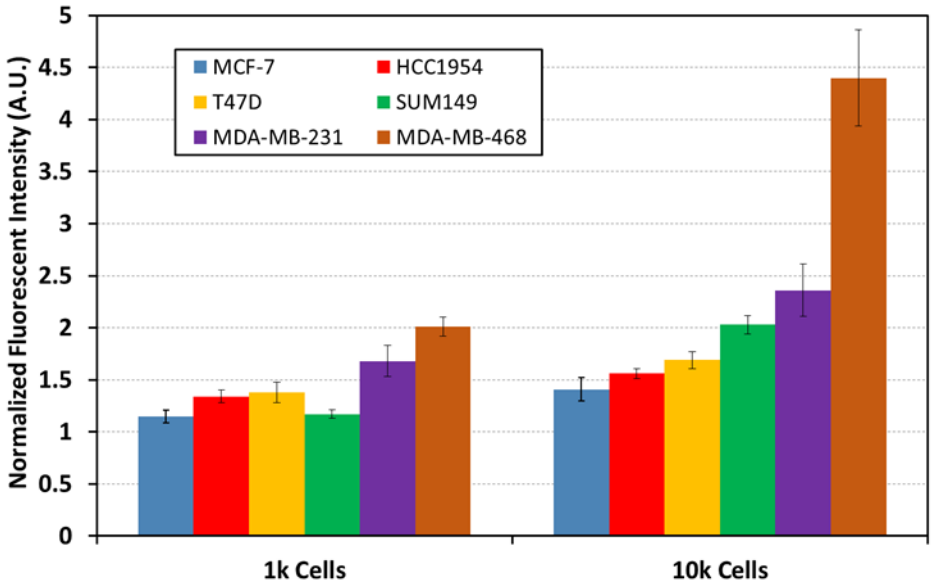

Figure 3-11 The fluorescent intensity of the substrate versus different number of cells and different cell lines (60 minutes reaction time). The fluorescent intensity was normalized to the well with no cell. The wells with more cells have higher proteolytic activity, and the MDA-MB-468 and MDA-MB-231 expresses higher proteolytic activity than other cell lines. (N = 6 wells in 96-well plate)

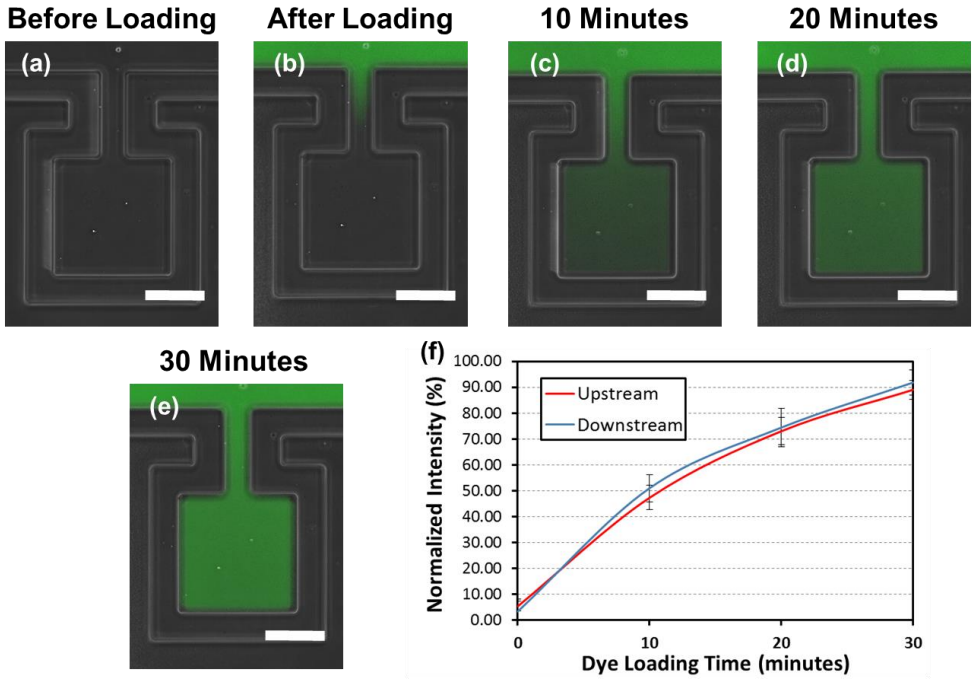

Figure 3-12 Diffusion of fluorescent substrate into the chamber. (a) Before loading of the substrate. (b-e) 0-30 minutes after loading the substrate into the main channel. (f) The fluorescent intensity inside the chamber versus the time. (N = 10 chambers) The increase

of the fluorescent intensity indicates the diffusion of the substrate into the chamber. The intensity in the chamber can reach 90% of that in the main channel within 30 minutes. (scale bar: 100 µm)

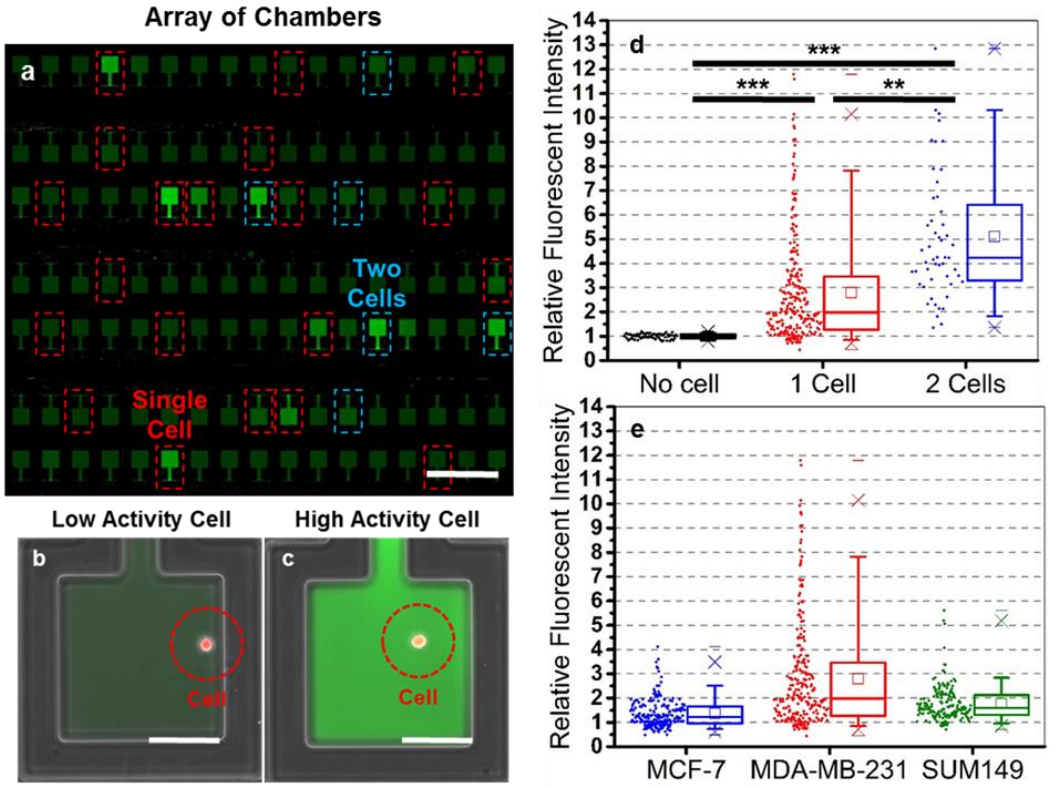

*Figure 3-13 Single cell proteolytic assay. (a) An array of cell chambers. The chambers circled with red dashed lines captured one cell, and the ones circled with blue dashed lines captured 2 cells. (scale bar: 1 mm) (b, c) Representative case of (b) a low proteolytic activity cell and (c) a high proteolytic activity cell. (scale bar: 100 µm) The fluorescent intensity reflects the proteolytic activity. (d) A box plot of the proteolytic activity of MDA-MB-231 cells; each dot represents one data point. The chambers containing 1 cell have significant higher fluorescent intensity than the empty chambers, and the chambers containing two cells are even brighter. (N ~ 200 cells for 0 and 1 cell per chamber, N ~ 50 for 2 cells per chamber) (e) A box plot of the proteolytic activity of MCF-7, MDA-MB-231, and SUM149 cells. A portion of MDA-MB-231 cells have very high activity, while the cells are more homogeneous for two other cell lines. (N ~ 200 cells for all three cell lines) * refers to P < 0.05, ** refers to P < 0.01, and *** refers to P < 0.001.*

For quantification of the proteolytic activity of single cells, we loaded single cells into the chambers as describe above. After that, the FRET substrate was loaded into the chamber and the chamber was isolated (Figure 3-12). Figure 3-13a-c shows an array of

chambers containing single cells and the enlarged view of cells with high and low proteolytic activity. To cancel the background noise, we normalize the fluorescent intensity using the empty chambers on the same chip to minimize the device to device variation. After normalization, the fluorescent intensity of chambers containing 1-2 cells are significantly higher than that of empty chambers (Figure 3-13d). When we characterizing three different cell lines: MCF-7, MDA-MB-231, and SUM149, different patterns of heterogeneity were observed (Figure 3-13e). The distribution of MDA-MB-231 has a long tail of highly active cells, while the SUM149 and MCF-7 cells are more uniform (Figure 3-14). Since bulk assays only report the average behavior within a population, obscuring the contribution of highly active sub-populations, no previous study of cancer cell protease activity has successfully observed these heterogeneous behavioral patterns.

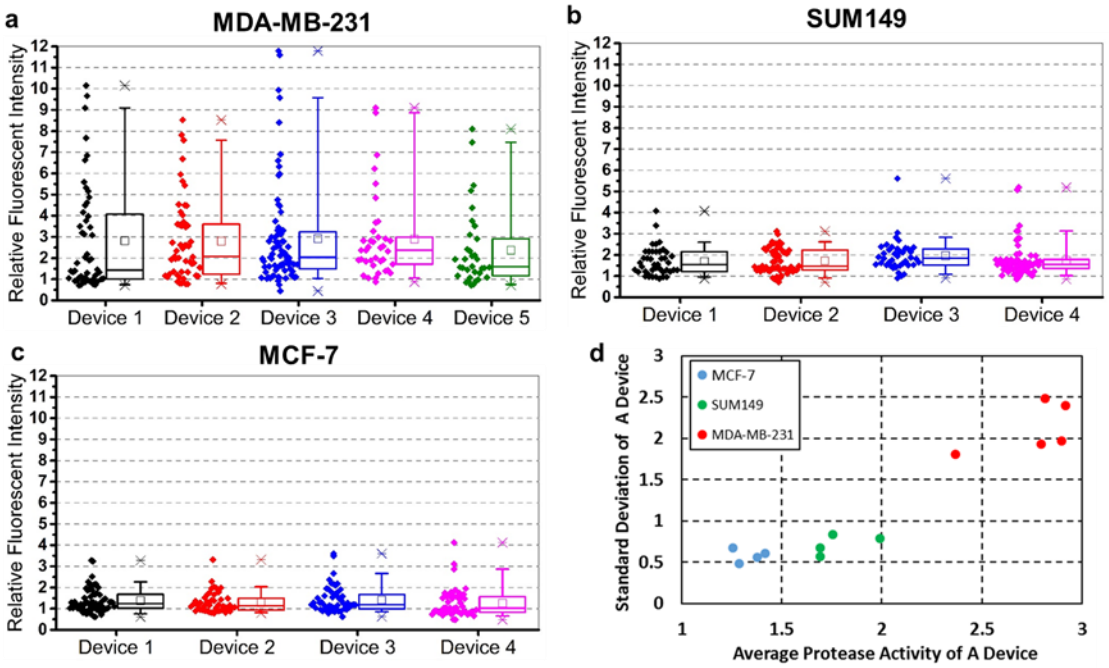

Figure 3-14 The assay variation between devices. The raw data of single cell proteolytic assays of (a) MDA-MB-231, (b) SUM149, and (c) MCF-7 cells. (N = 30-50 cells per device) (d) The distribution of devices: x-axis is the average of proteolytic activity, and the y axis is the standard deviation of cells in that device.

3.4 Intraclonal heterogeneity in proteolytic activity

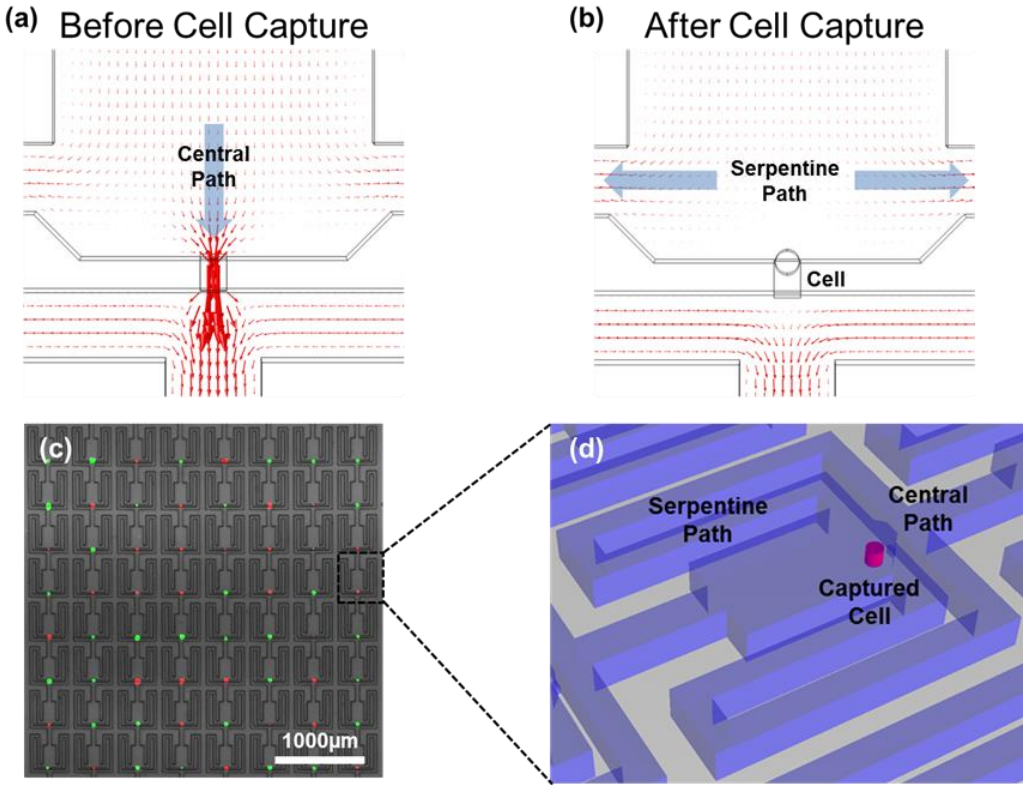

Figure 3-15 The hydrodynamic cell capture scheme. (a) The simulated (COMSOL 4.2) flow pattern before cell capture. The red arrows indicating flow direction and velocity suggest that the cells are likely to be guided to the capture site and get captured. (b) After cell capture, the captured cell blocks the flow, resulting in subsequent cells being guided into the serpentine path. (c) An 8x8 array of captured single cells. Two cell lines (red and green were different cell lines). were mixed for the demonstration. (d) Enlarged schematic of a cell capture chamber.

When quantifying the proteolytic activity of single cells in the previous experiment, we observed the innate heterogeneity expressed within a population. However, several recent studies have expounded the importance of microenvironmental regulation to cell heterogeneity [27–29]. A challenge then arises when trying to differentiate the contributions of the microenvironment from the innate heterogeneity of the cell population. To alleviate this complication, we can analyze microenvironmental effects on clonal populations. A microfluidic platform previously developed in our group, provides an ideal

solution by facilitating the creation of single cell-derived (clonal) spheres (Figure 3-15) [30]. The location of the cells within the sphere provides different microenvironmental cues that modulate protease activity of the clonally identical cells.

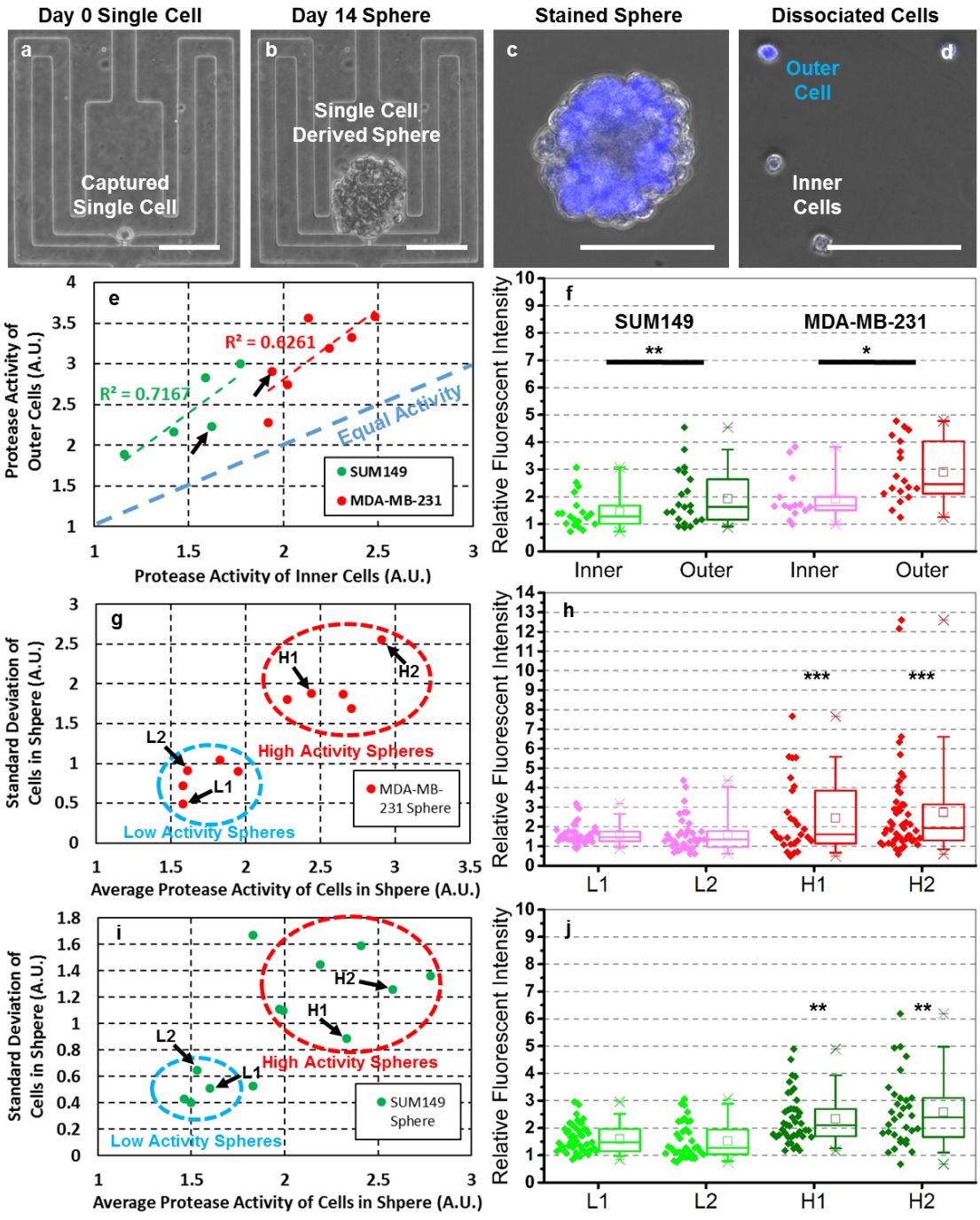

*Figure 3-16 Intraclonal and interclonal heterogeneity of proteolytic activity using single cell derived cancer spheres. (a - d) The process of MDA-MB-231 clonal sphere formation and analysis (scale bar: 100 µm): (a) cells captured on day 1 on non-adherent substrate, (b) single cells form sphere after 14 days, (c) sphere staining (blue cell tracker) to distinguish the outer and inner cells, and (d) the dissociation of spheres for single cell analysis. (e) The proteolytic activity of inner and outer cells, the x-axis is the average activity of inner cells and y-axis is the average activity of outer cells. Red and green dash line show the correlation between inner and outer cells by regression. The blue dash line is the equal activity line, showing that the average outer cells have higher activity than inner ones for all spheres. (N = 5 spheres or around 150 cells for SUM149, N = 7 spheres or around 200 cells for MDA-MB-231) (f) The box plot of proteolytic activity of inner and outer cells from representative SUM149 and MDA-MB-231 spheres (indicated in the previous figure by black arrows). (N ~ 20 cells for all conditions). (g) The average (x-axis) and standard deviation (y-axis) of proteolytic activity of MDA-MB-231 spheres. The spheres are clustered into two categories (high and low) based on protease activity. (N = 10 spheres). (h) The box plot of two representative MDA-MB-231 spheres from high and low activity categories (indicated in the previous figure by black arrows). (N ~ 40 cells for all spheres). (i) The average (x-axis) and standard deviation (y-axis) of proteolytic activity of SUM149 spheres. The spheres are clustered into two categories (high and low) based on protease activity. (N = 14 spheres). (h) The box plot of two representative SUM149 spheres from high and low activity categories (indicated in the previous figure by black arrows). (N ~ 40 cells for all spheres). * refers to P < 0.05, ** refers to P < 0.01, and *** refers to P < 0.001.*

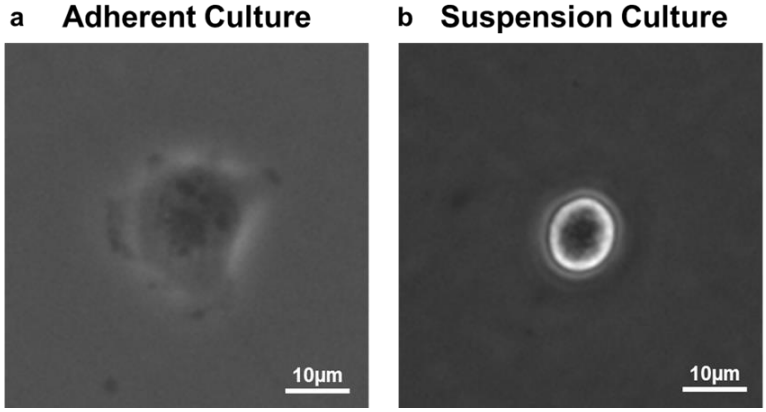

Figure 3-17 The comparison between (a) adherent culture of MDA-MB-231 on glass and (b) suspension culture of MDA-MB-231 on polyHEMA coated glass.

For single cell derived sphere formation, a cell suspension was loaded into the device and single cells were captured at the capture site (Figure 3-16a). As the bottom surface of the platform was coated with polyHEMA, a non-adherent polymer (Figure 3-17), the single cell grows in suspension to form a clonal sphere (Figure 3-16b and Figure 3-18). After 14

days, we dissembled the device to retrieve the spheres. Current research suggests that we should expect different protease activity at the edge of a tumor as compared to the center [29], as cells at the edge of the tumor must remodel the surrounding ECM in order facilitate further tumor growth and spread. In order to examine this phenomenon in a small scale model, we demonstrated the differences in protease activity between inner and outer cells in a harvested sphere. Cell tracker was introduced into the device to stain the spheres but was washed away after only 3 minutes of incubation. This short incubation time combined with the limited diffusion within the sphere resulted non-uniform staining of the cells in the sphere (i.e. cell fluorescence on the sphere periphery was considerably brighter than that in the center) (Figure 3-16c). Then, the sphere was dissociated into single cells by trypsin, and the difference in fluorescent intensity was preserved (Figure 3-16d) [31]. Using the presented method, we demonstrated the intraclonal heterogeneity (caused by microenvironmental effects) of two cell lines: MDA-MB-231 and SUM149 (Figure 3-16e). For all the spheres we tested, the average proteolytic activity of outer cells is higher than that of inner cells. In a representative sphere (Figure 3-16(f)), outer cells have wide dynamic range of activity, while the all inner cells only show low protease activity. Using a mixture of inner and outer cells from different single cell derived spheres leads to less significant results as the innate interclonal differences in proteolytic activity are comparable in magnitude (Figure 3-19). The results support that the internal cells in a tumor are inactive in proteolytic activity [2].

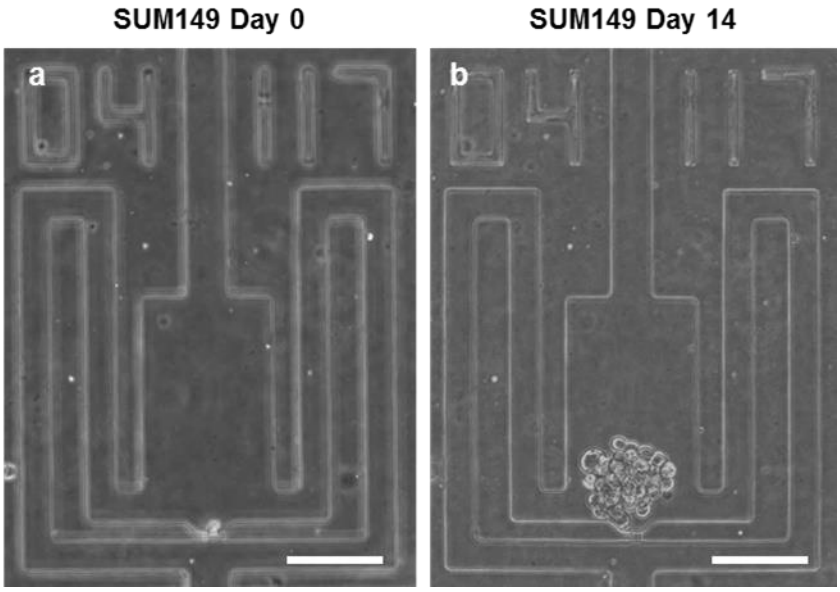

Figure 3-18 The sphere formation of SUM149 in the single cell suspension culture chip. (a) Right after cell capture, we have single cell captured. (b) 14 days after cell loading, the single cell grew to a sphere.

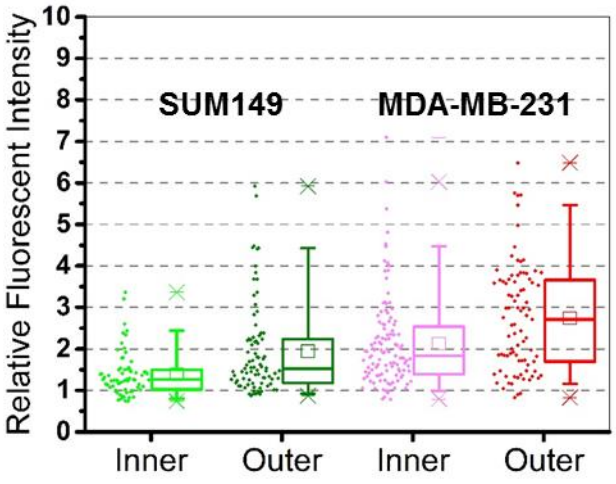

Figure 3-19 The comparison between inner and outer cells of all spheres (N ~ 100 for all conditions). Although the outer cells have higher proteolytic activity, the difference is less significant than comparing the differences observed between the spheres.

3.5 Interclonal heterogeneity in proteolytic activity

Although we observed intraclonal heterogeneity caused by microenvironmental differences, the activities of inner and outer cells from the same sphere are highly correlated

as well (Figure 3-16e). To further examine this phenomena, we characterized more single cell derived spheres of MDA-MB-231 and SUM149 cells (Figure 3-20 and Figure 3-21). When we mapped the MDA-MB-231 spheres on a 2D plot using the average activity as the x-axis and standard deviation within sphere as the y-axis, two subtypes of spheres clustered: those with low activity and variation and those with both higher average activity and variation (Figure 3-16g). The single cell activities of four representative MDA-MB-231 spheres are shown in Fig. 3h. SUM149 spheres follow similar trend (Figure 3-16i, j). As we analyzed cells from different spheres using different devices, it was also important to examine whether the variation was caused by the device-to-device variation. We found that the devices loaded with bulk cells have significantly lower variation than those loaded clonal cells (Figure 3-14), indicating the heterogeneity observed was likely a result of the sphere culture and not the devices themselves. In addition, we plotted the sphere size (related to proliferation rate) versus the proteolytic activity of its dissociated cells (Figure 3-22). Surprisingly, we found that though larger spheres have more cells, the size of sphere has no correlation with single cell proteolytic activity. These results demonstrate the capability of presented approach to dissect the intraclonal and interclonal heterogeneity of cell lines.

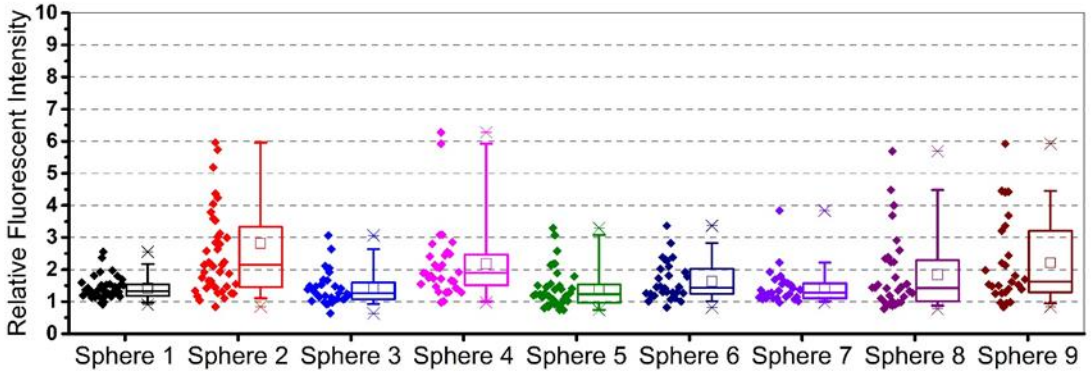

Figure 3-20. The raw data of single cell proteolytic activity from MDA-MB-231 spheres. (N = 30-50 cells per sphere)

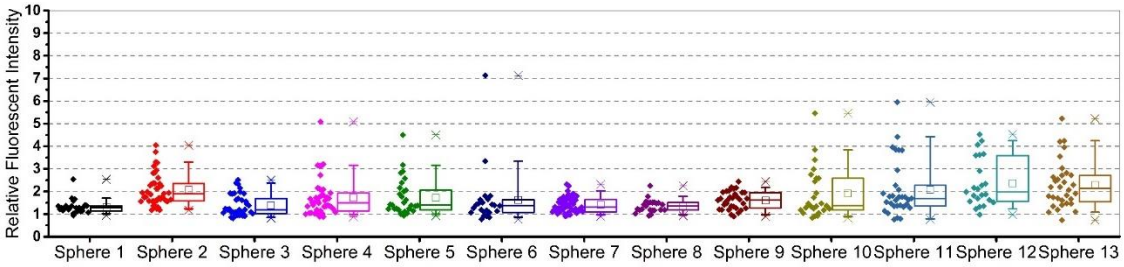

Figure 3-21 The raw data of single cell proteolytic activity from SUM149 spheres. (N = 30-50 cells per sphere)

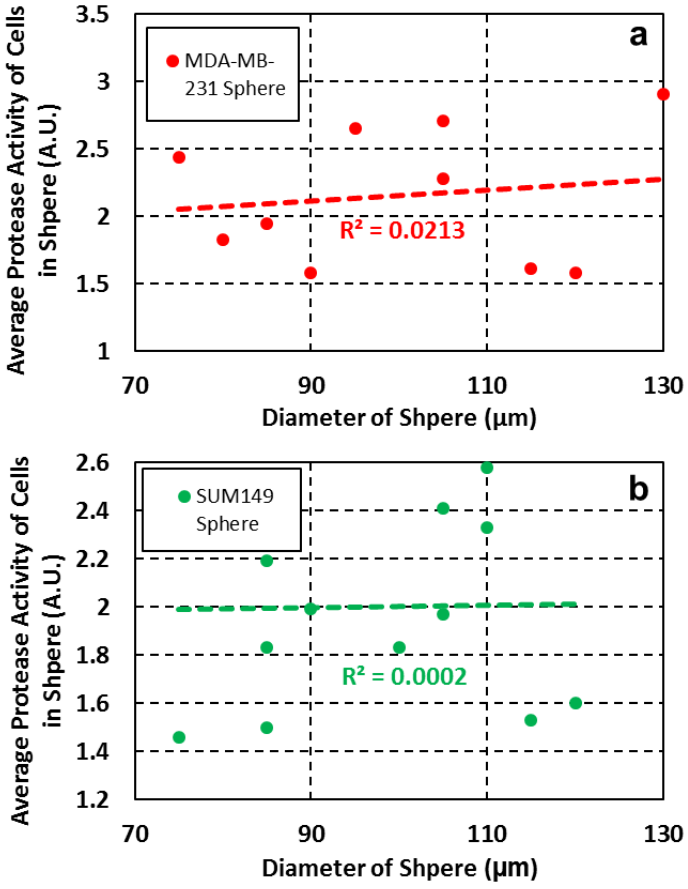

Figure 3-22 The correlation between the size of (a) MDA-MB-231 (N = 10 spheres) and (b) SUM149 (N = 14 spheres) spheres, and the proteolytic activity of single cells in that sphere. No correlation was observed.

3.6 Dynamics of proteolytic activity

After characterizing the heterogeneity between cells, we next investigated the variation of proteolytic activity of the same cell at different time points (dynamics). As the time constant of protein translation is several hours [10], we measured single cell protease activity every 8 hours for 2 days. These time points provide sufficient time for dye diffusion and cell recovery between assays as well (Figure 3-23). Fig.4a-f show the activity of the individual cells at different time points, and the cells were clustered based on dynamics in the heat maps (Figure 3-24g, h). Different dynamics between MDA-MB-231 and SUM149 were observed. The MDA-MB-231 had wide dynamic ranges of activity and sharp short pulse of high activity, causing high variation between time points (Figure 3-24i). We were concerned that the sharp changes were the result of cell divisions, but when cells were monitored for division events, the data show no difference of activity before and after cell division takes place (Figure 3-25). Compared to the activity of MDA-MB-231, the activity of SUM149 cells changes gradually and slowly, resulting in low variation between time points (Figure 3-24j). The representative cases of two cell line are plotted in Figure 3-24(k, l). In addition to demonstrate the cell stability on-chip, we verified that the average proteolytic activity of the single cell population remained stable at all time points (Figure 3-26).

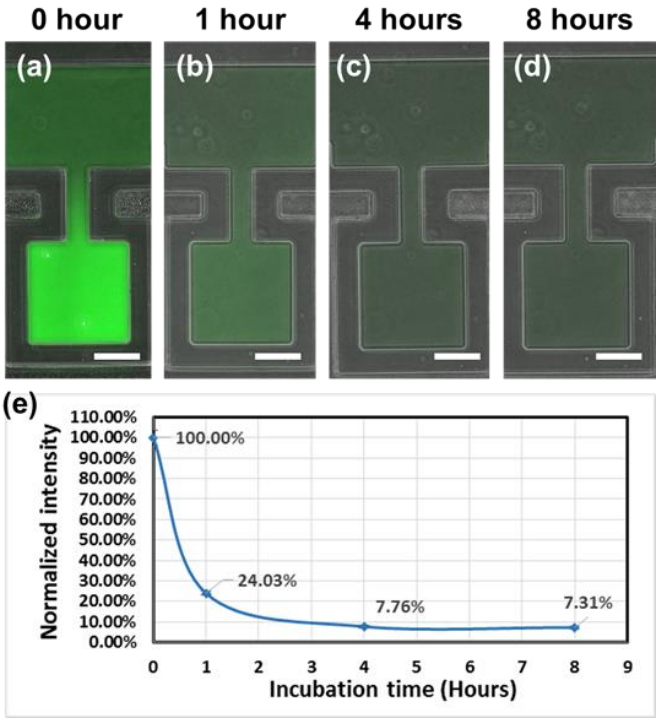

Figure 3-23 Diffusion of fluorescent substrate out of the chamber. (a) Chambers with substrate right after refilling culture media in the main channel. (b-d) 1, 4 and 8 hours after the media refill. (e) The fluorescent intensity inside the chamber versus the time. The decrease of the fluorescent intensity shows the diffusion of the substrate out of the chamber. The intensity in the chamber can be reduced to below 10% for the next proteolytic assay after 4 hours (N = 12 chambers). (scale bar: 100 µm)

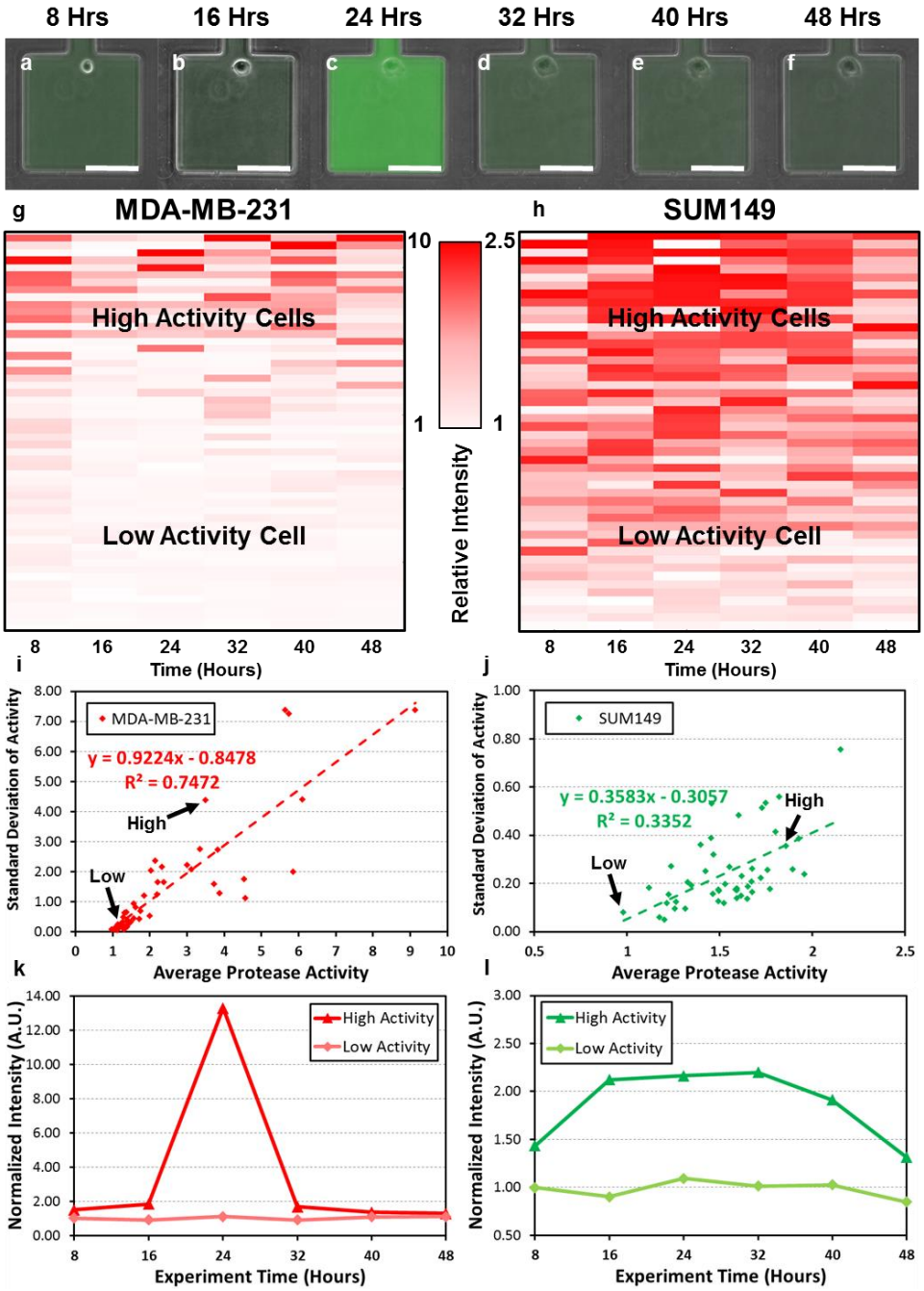

Figure 3-24 Dynamics of single cell proteolytic activity. (a-f) The dynamic activity of a MDA-MB-231 cell: (a) 8 hours after loading, (b) 16 hours, (c) 24 hours, (d) 32 hours, (e) 40 hours, and (f) 48 hours. (scale bar: 100 µm) (g) The dynamics of the MDA-MB-231 cells plotted in a heat map. Red color represents high activity (relative fluorescent intensity: 10), and white color represents low activity (relative fluorescent intensity: 1). The cells were sorted based on the summation of the average and the standard deviation of their proteolytic activity. (N ~ 50 cells for all 6 time points). (h) The dynamics of the SUM149 cells plotted in a heat map. Red color represents high activity (relative fluorescent intensity: 2.5), and white color represents low activity (relative fluorescent intensity: 1). The cells

were sorted based on the summation of the average and the standard deviation of their proteolytic activity. (N ~ 50 cells for all 6 time points). (i) The average (x-axis) and standard deviation (y-axis) of proteolytic activity of MDA-MB-231 cells at different time points. Red dashed line shows the regression of the correlation between average and standard deviation. (j) The average (x-axis) and standard deviation (y-axis) of proteolytic activity of SUM149 cells at different time points. Green dash line shows the regression of the correlation between average and standard deviation. (k) Example dynamics of representative high activity and low activity MDA-MB-231 cells (indicated in (i) by black arrows). (l) Example dynamics of representative high activity and low activity SUM149 cells (indicated in (j) by black arrows).

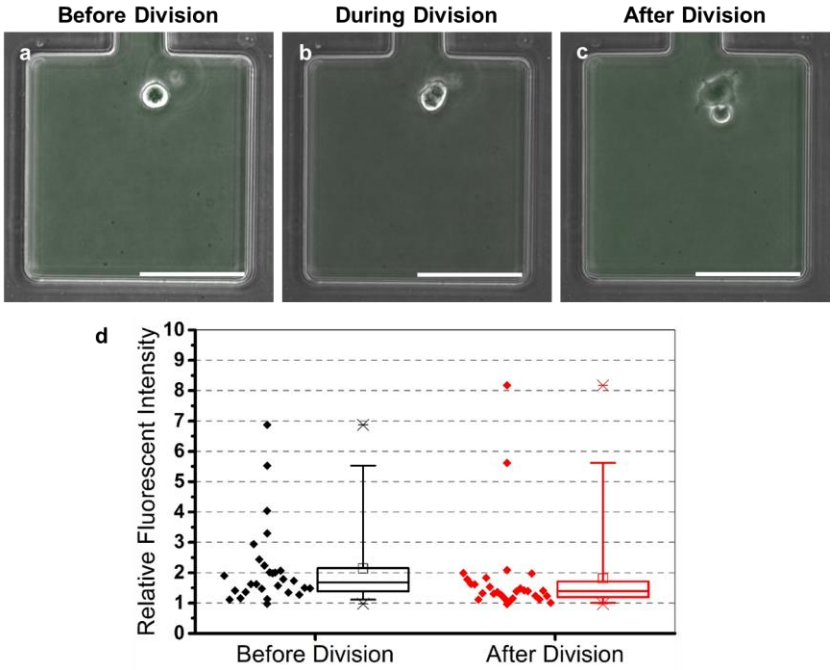

Figure 3-25 Cell division and proteolytic activity: (a) cell before division, (b) cell during division, (c) cell after division, and (d) the proteolytic activity of cell before and after division. (N = 20-30 cells per condition)

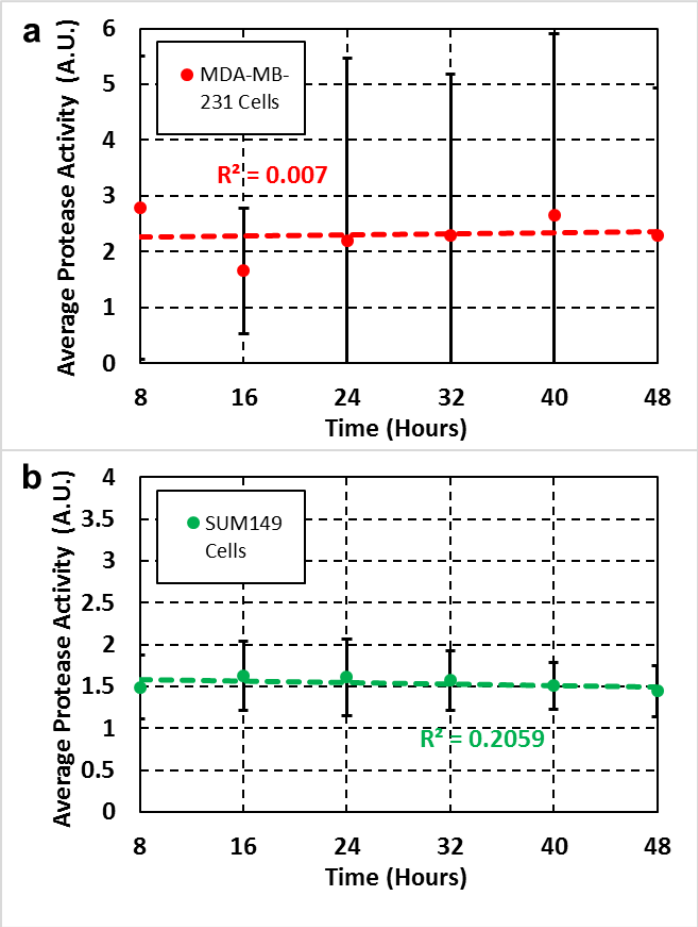

Figure 3-26 Average proteolytic activity of cells at different time points: (a) MDA-MB-231 and (b) SUM149 cells. Low correlation between time and activity indicates the cell behaviors didn't change or degrade over time. (N ~ 50 cells)

3.7 Chapter Summary

The regulation of proteases is a key issue in cancer metastasis. Although proteolytic assays can be performed in a conventional (dish based) manner, the presented single cell approach provides three advantages over conventional approaches: (1) the capability to handle small samples (~10-100 cells), (2) the potential to investigate intraclonal and interclonal heterogeneity, and (3) ability to monitor and track individual cells over time. Conventional approaches need hundreds of thousands of cells for an assay. A patient biopsy sample may be barely enough for a single trial. Circulating tumor cell (CTC)

samples, which are less invasive to acquire from the patients, typically have only 10-100 cells from 7.5mL of patient blood. It is unfeasible to use dish based approaches for the assay of these CTCs or other primary samples, so there is an outstanding need for the capability to handle small samples. Though hydrodynamic cell capture scheme can elegantly position single cells and achieve high (~80-90%) single cell capture rates (number of captured cells/number of chambers), the cell capture efficiency (number of captured cells/number of cells used) is typically lower (<10%). Two fundamental limitations are: (1) cells are lost in the dead volume that never flows into the device and (2) continuous flow in the hydrodynamic scheme inevitably carries a portion of the cells to the outlet without the cells being captured. The presented platform drives all the sample into the chambers to achieve a negligible dead volume, and all loaded cells will be retained in the chambers. As a result, we can achieve a high (>50%) capture efficiency, even for small samples. Compared to the approaches using open micro-well array [32], the presented approach has advantages that include (1) easy and robust media/reagent exchange, (2) enclosed microfluidics to minimize the evaporation, and (3) reliable isolation between chambers. As such, this single cell enzymatic assay platform specializes in the single cell analysis of small samples, something not possible in previous approaches.

Using the platform, we can easily investigate the cellular heterogeneity that often complicates cancer treatment and analysis. Previous studies have used mixtures of different cell lines to mimic the in-vivo heterogeneity and showed that heterogeneity in and of itself contributes to invasion. However, in these experiments, we demonstrated these differences exist within single lines. MDA-MB-231 cells show a wide distribution of proteolytic activity, while SUM149 are more homogeneous and less active. The high proteolytic

activity and their innate heterogeneity may contribute to these cells stronger metastatic potential as compared to SUM149.

As suggested in the literature, microenvironment had a large effect on proteolytic activity. Sphere formation provides one approach to look at these effects as the location of the cells within a sphere provides different cues that modulate protease activity. By beginning with single cell derived spheres, the differences in activity can be attributed to these cues alone. As we have a limited number of clonal cells from a sphere, it is necessary to be able to analyze small samples. First, we found that within a 3D sphere, the inner cells have lower proteolytic activity, while the outer cells are highly active. This result verifies the heterogeneous nature of proteolytic activity in a tumor sphere and supports the importance of microenvironmental effect during tumor growth [29]. Located at different positions within a sphere, the descendant cells from the same progenitor can have significantly different proteolytic activity. Although the cells in the same sphere are heterogeneous, the proteolytic activities of inner cells and outer cells are still correlated. This heterogeneity was observed despite the fact that cancer cell lines are believed to be more homogeneous than primary patent samples. For both MDA-MB-231 and SUM149 cells, the clones can be classified as low activity or high activity spheres. The cells in high activity clones are highly heterogeneous, while those in low activity clones have homogeneously low activity. We also tried to correlate the sphere size, which is proportional to the cell proliferation rate, and the proteolytic activity. Surprisingly, we didn't observe correlation between the size and the average proteolytic activity, indicating that the regulation of proteolytic activity seems to be independent of the cell proliferation and cell cycle. As most cancer cell heterogeneity studies are based on bio-markers using

fluorescence-activated cell sorting (FACS), the next step to integrate our findings into the literature more thoroughly would be to investigate the relationship between markers and the proteolytic activity. One possible study would be to investigate the relationship between protease activity and cancer cell "stem-ness". In breast cancer, researchers have identified a rare subpopulation of tumorigenic cancer stem-like cells (CSC), often thought to be the source of drug resistance and tumor metastasis [33–35]. These populations are typically identified by ALDH or CD44+/CD24- [34,35]. As proteases are required for cancer cell invasion, the correlation may clarify the roles of different sub-populations in tumor metastasis.

Though we successfully investigated inter and intra- clonal heterogeneity, it was still not known whether high proteolytic activity is the intrinsic characteristic of a particular cell or if the proteolytic activity of single cell follows a stochastic pattern, changing with time. Thus, we monitored the proteolytic activity of the same cell every 8 hours to understand its dynamics. The preliminary data suggest that the proteolytic activity of MDA-MB-231 seems to follow a stochastic pattern. The same cell can have wide (10 times) dynamic range of proteolytic activity at different time points, and the activity can increase and decrease rapidly within 8 hours. We first suspected the correlation between cell cycle and the proteolytic activity; however, when we examined the chambers with proliferating cells, we found that the proteolytic activity remains stable before and after cell division. Also, no difference can be found between proliferating and non-proliferating cells. This supports that the previous finding that proliferation rate seems to be independent of its proteolytic activity. Compared to MDA-MB-231, the proteolytic activity of SUM149 changes slowly. The elevated activity is maintained for 24 hours or longer, so the standard

deviation of activity at different time points is significantly lower than that of MDA-MB-231 cells. This difference indicates that the protease regulating mechanisms of two cell lines are possibly quite different. These types of valuable information can be easily hidden when averaging activities of many cells are examined using dish based approaches. This data shows that the single cell analysis approached present here provides a novel tool to unveil the enzymatic activity pattern in time domain, enabling more mechanistic studies as the next step.

Chapter 4
HYDRO-SEQ: CONTAMINATION-FREE HIGH-THROUGHPUT SINGLE-CELL RNA-SEQUENCING FOR RARE CELL POPULAITONS

Although massively parallel single-cell RNA-sequencing can characterize cellular heterogeneity at scale, it remains challenging to analyze contaminated rare samples such as circulating tumor cells (CTCs). Here, we present Hydro-Seq, a scalable hydrodynamic bead-cell-pairing technique with high cell capture efficiency, high-throughput, and contamination removal capability. We successfully achieved whole transcriptome sequencing of 666 CTCs from 21 breast cancer patient samples, identifying cellular heterogeneity in critical biomarkers of tumor metastasis and therapy.

4.1 Introduction

Recent advances in single-cell RNA-sequencing (scRNA-seq) have enabled high-throughput analysis of cellular heterogeneity and identification of cellular types by their gene signatures [1-3]. By pairing single barcoded beads with single cells in droplets or micro-wells, mRNA molecules from a single cell can be uniquely labelled by a barcode and identified using single-cell whole transcriptome analysis [1-4]. However, the low efficiency of droplet-based technologies such as Drop-seq as well as the presence of contaminating cell populations have limited the applicability of these techniques in analysis of rare cell populations such as circulating tumor cells [1]. Due to the inefficiency of bead-cell pairing, thousands of cells are required to achieve accurate cell readout. As a result, it remains challenging to apply the existing scRNA-seq technologies to analyze rare cell

populations such as circulating tumor cells in cancer patients, which contain limited cells of interest (10-100 cells) and massive contamination such as debris, cell-free nucleic acids, and background cells (erythrocytes and leukocytes). Alternatively, single-cell picking methods including capillary suction and dielectrophoretic microfluidics have been used to isolate a few cells of interest labelled by fluorescent staining [4-6]. Nevertheless, these techniques are constrained by their limited throughput and reliance upon antibody-based marker expression. Commonly utilized markers such as EpCAM and cytokeretins may miss a substantial population of CTCs which fail to express these proteins [5, 6]. Thus, the development of a high-throughput technology capable of efficiently capturing and molecularly interrogating CTCs at single-cell resolution would have significant clinical utility including its use in treatment selection and therapeutic monitoring [2, 5].

4.2 Hydro-seq design

Here, we present Hydro-Seq, a high-efficiency-cell-capture scRNA-seq platform for profiling rare cell populations such as CTCs. (Figure 4-1 and Figure 4-2). Hydro-Seq represents three major technological advances: (1) size-based high-efficiency single-cell capture, (2) chamber washing capability for cellular and acellular contaminant removal, and (3) scalable array design for massively parallel analysis. The massively parallel filter structure enables high-efficiency cell capture of greater than 90%. We engineered chips consisting of 800 chambers per chip capable of accommodating 1-100 CTCs from 10mL of patient blood, a design that can be expanded to 12,800 chambers for other applications [5, 7]. In this design, there are 16 branch channels with each channel containing 50 cell-capture chambers. Each chamber contains ~1nL volume, similar to the well and droplet

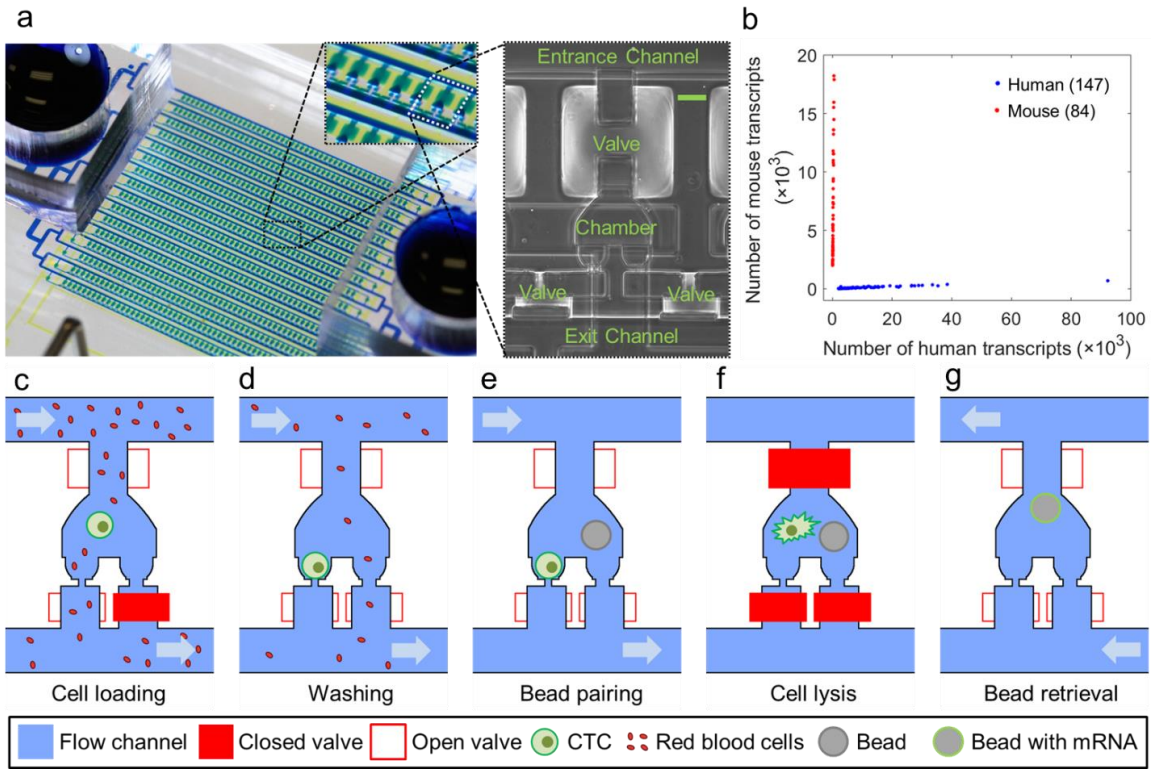

Figure 4-1 Hydro-Seq, a high capture efficiency scRNA-seq platform for contaminated rare samples. (a) Hydro-Seq bead cell pairing chamber array for hydrodynamic cell capture and washing. Each chamber contains a cell capture site and a bead capture site, each controlled by a valve for flow control. (Scale bar: 50μm) (b) Mixed species analysis highlights the single-cell resolution RNA-sequencing with zero cellular cross contamination. Each red dot represents a mouse cell and each blue dot represents a human cell. (c-g) Hydro-Seq operation flow. (c) Bead capture valve is closed during sample loading. The smaller red blood cells flow through the capture until a larger cell (larger than 12μm) blocks the channel for cell capture. (d) After cell capture, the bead capture valve is then opened to wash the contaminations away. (e) After removing contaminations in the chamber, the bead is then loaded for pairing. (f) Lysis buffer introduced into the chamber while cell remains in the dead volume. After closing all the valves, the cell is pushed to the middle of the chamber for lysis and mRNA capture. (g) By opening all the valves and introducing a back flow, the beads can be retrieved to the inlet for downstream preparation.

volume reported in scRNA-seq methods [1,2]. For bead-cell pairing, each chamber contains one cell capture site and one bead capture site each of which can be blocked by different pneumatic valves (Figure 4-1b,e). To enable cell lysis and other washing processes, two washing channels with valve controls were added to the entrance and the exit of each branch channel (Figure 4-1c,d).

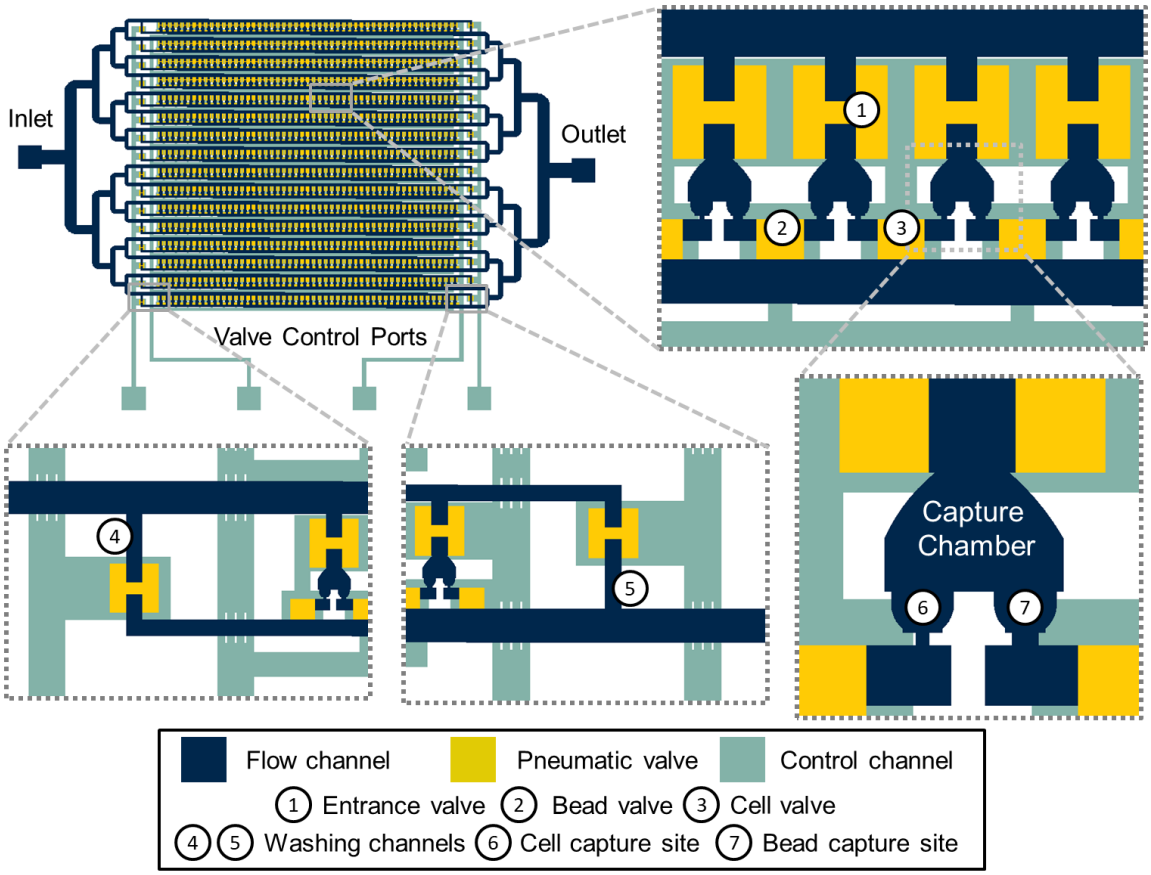

Figure 4-2 Design of Hydro-Seq technology. Integrated microfluidic circuit design with valve controls for high-efficiency cell capture and contamination removal. (a) For CTC application, Hydro-Seq chip is designed with 16 identical branch channels in parallel. Each branch channel consists of 50 chambers for bead cell pairing, totaling 800 chambers per chip. (b) A closer view highlights the arrangement of parallel chambers. To minimize area consumption, each chamber shares the valve with its neighboring chamber. The entrance valve has a peak height of 45μm and area of 200μm×200μm to enable bead and cell loading. The cell and bead valves in height (15μm) and area (100μm×100μm) for high-density chamber arrangement. (c-d) At the upstream and downstream of the branch channel, valve-controlled wash channels enable channel washing during sample loading and delivery of lysis buffer to the entrance channel for lysis. (e) A closer view highlights the pairing chamber with cell and bead capture sites.

4.3 Device fabrication

The devices were made using soft-lithography fabrication process. The multi-layer layout of the chip was designed using AutoCAD 2016 (Autodesk®). The masks for

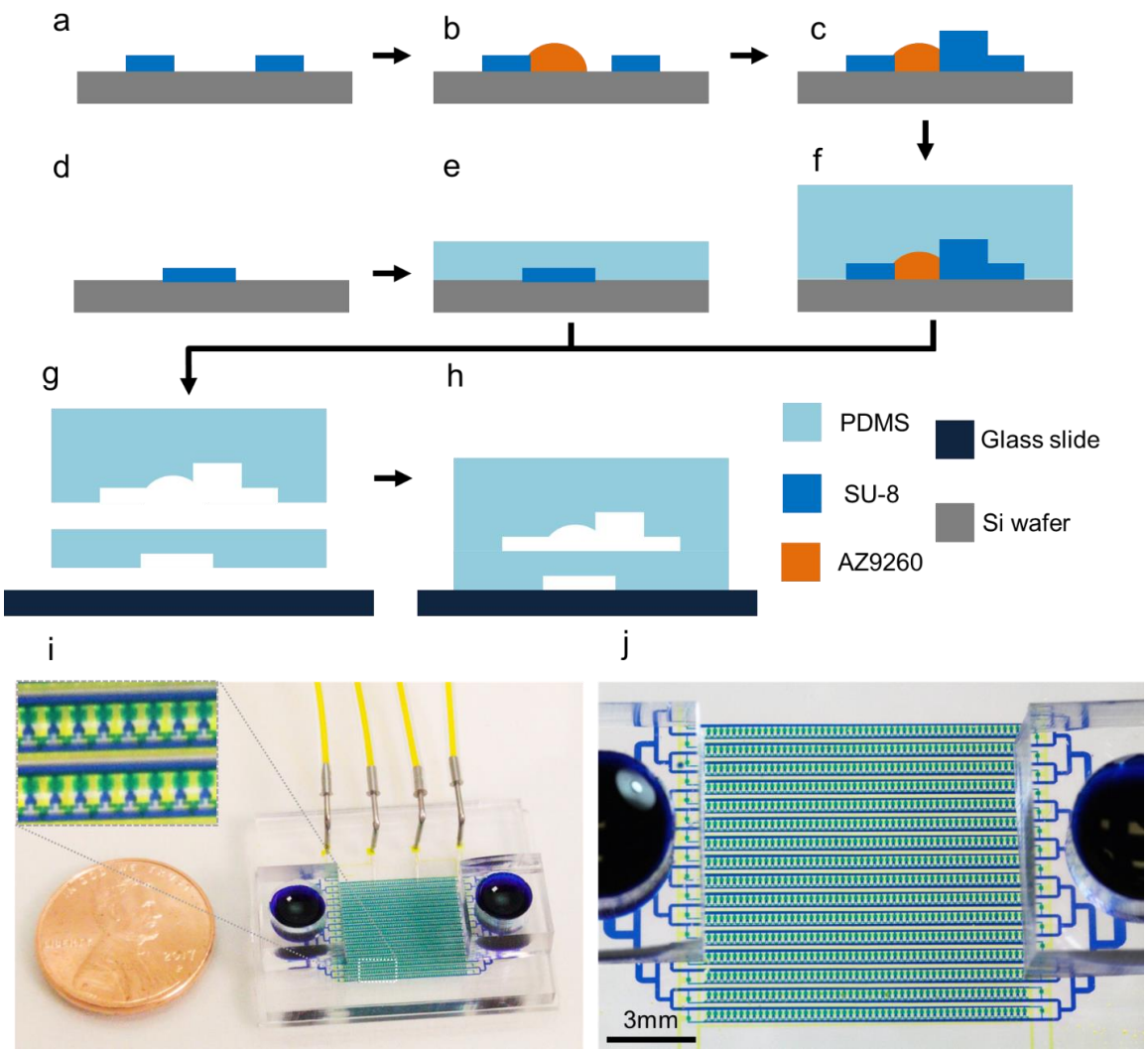

Figure 4-3 Multilayer fabrication of Hydro-Seq. (a) SU-8 patterned on the silicon wafer to create the mold for soft lithography fabrication. (b) With photoresist AZ-9260 patterning, the thermal flow process created curved structure for valve sealing. (c) The flow channel mold was made with 2 layers of AZ-9260 patterning and 4 layers of SU-8. (d) The valve control channel mold was made by one layer of SU-8 patterning. (e-f) After multilayer fabrication, self-assembled monolayer of silane coating is then applied to facilitate PDMS peeling. (e) The valve control layer made by spinning a 30µm PDMS. (f) The flow layer made by pouring 2mm PDMS on the mold for curing. (g) To assemble the layers together, the flow layer was detached from the mold. After alignment and surface plasma activation, the flow layer was bonded to the control channel. (h) The fabrication completed by attaching the PDMS layers to a glass substrate. (i) Photograph of the fabricated device with a US penny. Four valve control channels are connected to the device for valve manipulation. (j) Photograph of the high-density chamber array.

photolithography were made using a mask making instrument (µPG 101, Heidelberg instruments). The mold for the flow channel was fabricated with 10 µm, 20 µm, 40 µm,

and 100 µm thick SU-8 (Microchem) following the manufacturer's protocol. The valves were created using AZ®9260 (AZ Electronic Materials) with peak thickness of 15 µm and 45 µm after thermal reflow. The mold for the control channels was fabricated with 20 µm SU-8. The SU-8 mold was treated by vaporized Trichloro(1H,1H,2H,2H-perfluorooctyl) silane (448931 ALDRICH) under vacuum overnight to promote the release of cured PDMS. After coating, the mold was heated at 150 ºC on a hot plate for 10 minutes. PDMS (Sylgard 184, Dow Corning) was prepared by mixing with 10 (elastomer): 1(curing agent) (w/w) ratio, poured on flow channel molds, and cured at 85ºC overnight before peeling. Thin film of PDMS was spun onto the control channel mold and cured at 85ºC for one hour. After peeling the PDMS from the flow control mold, the PDMS piece and the thin film PDMS were treated using oxygen plasma (80W for 60 seconds) and bonded using MJB3 aligner (Karl Suss). The devices after bonding were heated at 80ºC overnight to ensure bonding quality.

4.4 Sample loading operation

Before loading CTCs into Hydro-Seq, we performed size-based CTC enrichment and the CTC enriched sample was then transferred to Hydro-Seq for scRNA-seq preparation[6, 8]. Although the enrichment step removes several orders of magnitude of blood cells, there are still significant numbers of residual erythrocytes and leukocytes in the enriched sample, which were further eliminated by Hydro-Seq (Figure 4-1). During cell loading, the bead flow channels and washing channels were blocked, so CTCs could be captured at the cell capture sites. As CTCs are typically larger than other blood cells, the cell capture site was designed as a channel with 10µm×10µm opening [5]. This channel size allows smaller leukocytes, erythrocytes, and platelets to pass through, while larger cells (CTC or larger

leukocytes) can be captured (Figure 4-1 and Figure 4-2). When a CTC is captured, it blocks the capture site and the following cells are diverted to other downstream chambers, ensuring that only a single CTC occupies each chamber. After cell loading, the bead capture site was then opened to wash away any contaminants remaining in the chamber. The barcoded beads were then loaded into the chip to pair with the captured cells. Since the beads have an average diameter of 40μm, the bead capture site was designed with a 20μm×25μm opening. Cell lysis buffer was introduced into the chambers to lyse cells, and the released mRNA hybridizes with the barcoded bead. Finally, the barcoded beads were retrieved with a back flow for downstream sequencing procedures, including reverse transcription, amplification, library preparation and paired-end sequencing. With the

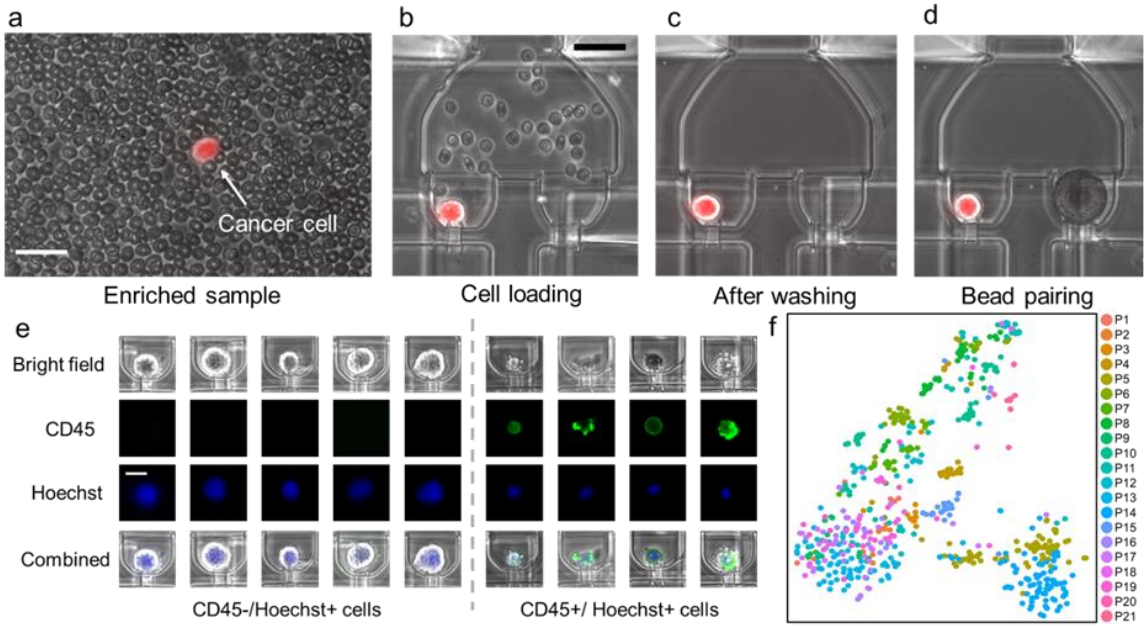

Figure 4-4 CTC loading in Hydro-Seq for immunostaining and single-cell sequencing. (a) After CTC enrichment, scRNA-seq of the CTCs is still challenged by the presence of many background blood cells. (Scale bar: 25μm) (b) Erythrocytes flowing through the chamber during sample loading. (Scale bar: 25μm) (c) By opening the bead valve, blood cells can be removed through the bead capture flow channel, achieving contamination-free single-cell isolation for bead pairing. (d) Pairing of the bead to a single cell for scRNA-seq. (e) With CD45 and Hoechst staining, larger CD45 positive nucleated cells were identified as leukocytes. (Scale bar: 15μm) (f) tSNE plot of all patient samples processed by Hydro-Seq. (666 CTCs from 21 patient samples with each color representing one sample.)

precisely controlled hydrodynamic capture operation, single CTCs can be selectively captured and paired with barcoded beads for scRNA-seq with high purity, high efficiency, and high throughput.

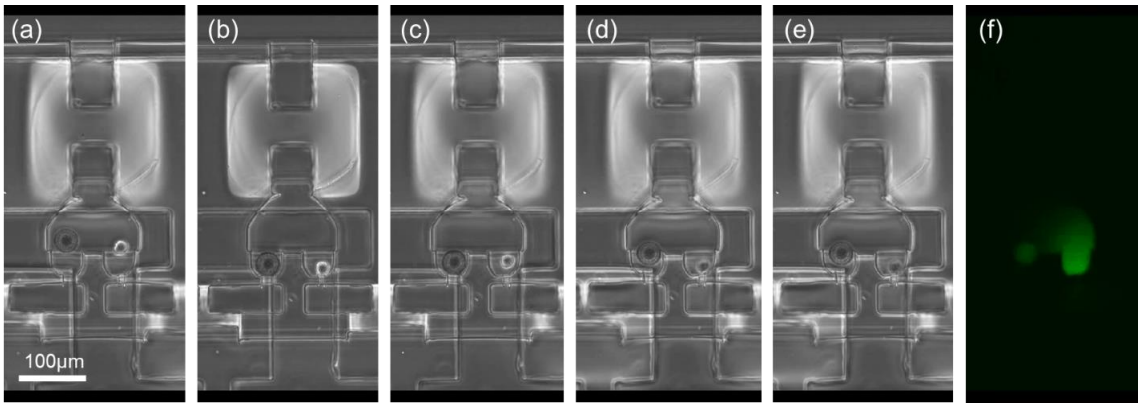

Figure 4-5 Cell lysis in a Hydro-Seq chamber. (a) Before lysis, the cell and bead are paired in the chamber with the valves closed. The lysis buffer is loaded to the branch channel. (b) After opening the valve, the flow pushes the bead and cell back to the capture site, and the lysis buffer flows into the chamber. As the cell seals the capture channel, the cell remains intact in the dead volume in the pocket. (c) After closing the valve, the operation creates flow in the chamber, exposing the cell to the lysis buffer. (d-e) The cell is lysed within 30 seconds. (f) The cell with green fluorescent protein (GFP) releases the cellular content, with the GFP highlighted by fluorescent imaging. (MDA-MB-231 GFP cells used in the experiment.)

4.5 Patient sample processing protocol

Whole blood from patients with metastatic breast cancer was obtained as part of an Institutional Review Board approved protocol (HUM00070190). All participants are over 18 years of age. All subjects were consented by the study team or research nurse prior to the scheduled blood draw using standard procedures for clinical research from the University of Michigan Comprehensive Cancer Center. For experiments with Labyrinth purified samples, the blood samples (~10mL) were processed using the protocol reported in a prior literature6. For experiments with Celsee purified samples, the blood samples

(~10mL) were processed by the Celsee PREP100 system following manufacturer's protocol. For both technologies, the samples were spun down to 100µL volume after CTC enrichment. After device priming and preparation, the CTC suspension was loaded to the device by inserting the pipette tip filled with the 100µL solution to the inlet. After closing all the wash channels and bead valves, the CTC suspension was loaded to the chip with a flow rate of 10µL/min driven by a syringe pump. After emptying the pipette tip, the tip was removed and 100µL PBS was added to the inlet. After washing with PBS at 10µL/min for 2 minutes, the bead valves were opened, and the flow rate was increased to 50µL/min to wash away residual red blood cells in the chamber. The PBS solution was refilled during the wash process. After washing for 3 minutes, the flow was stopped, and bead valves were closed again. To further remove contaminating cells in the chamber, 100µL PBS was added

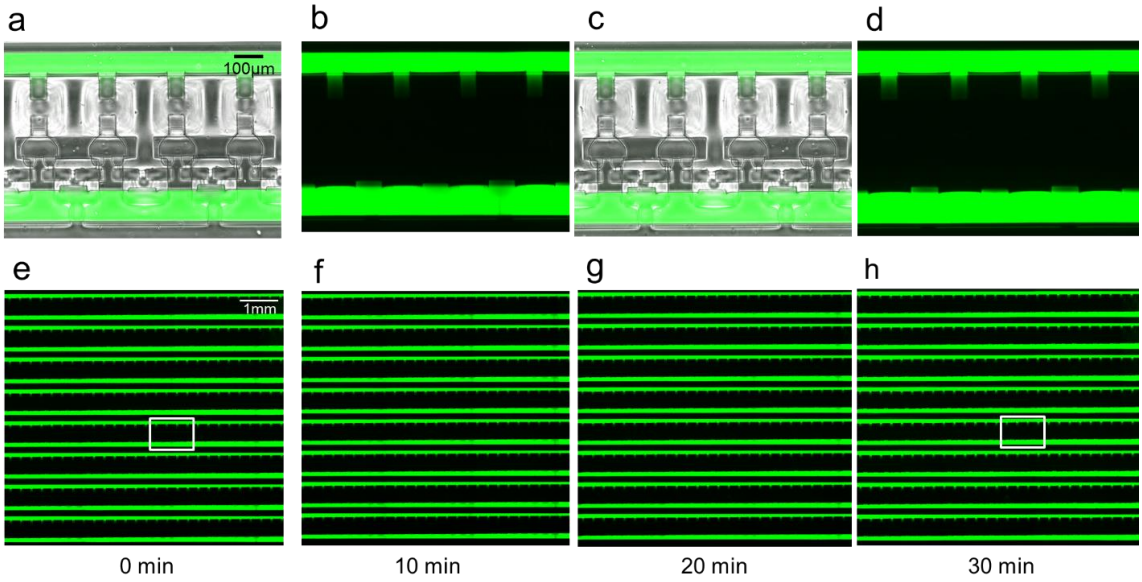

Figure 4-6 To ensure high quality single-cell analysis and prevent cross contamination between chambers, leakage test of the valve system was applied for 30 minutes, the time for mRNA incubation and on-bead capture. (a) Microscopic image showing the parallel chambers isolated by activating the pneumatic valves at 0 minute after introducing the fluorescent dye. (b) The corresponding fluorescent channel of image (a). (c) Microscopic image of the parallel chambers at 30 minutes after introducing the fluorescent dye. (d) The corresponding fluorescent channel of image (c). (e-h) Images highlight chamber isolation at scale at 0 minute (e), 10 minutes (f), 20 minutes (g), and 30 minutes (h), after introducing fluorescent dye. The white box in (e) is shown in (a), and the box in (h) is shown in (c).

to the outlet and a pipette tip was inserted to the inlet to retrieve the solution using a pipette. Then, the solution was loaded back to the chip and washed again using the same protocol in the first loading. After loading CTCs, the beads were loaded to the chip and prepared following the same procedure described in the mixed species session.

4.6 Cell capture efficiency

To demonstrate high cell capture efficiency of Hydro-Seq, we performed cell loading tests with ~100 cancer cells spiked into blood samples (from healthy donors) enriched by Celsee systems and achieved 90.43±6.08 % capture efficiency. After cell capture, a washing procedure was applied to remove contaminating cells and 89.70±5.06% of the captured cells remained. Finally, 89.60±6.39% of the remaining cells were successfully paired with a single bead for barcoding mRNAs. Accounting for the losses from cell loading, washing, and bead pairing, 72.85±2.64% of the initially loaded cell populations were successfully paired with a barcoded bead on chip.

4.7 Species-mixing experiment

To assess the fidelity of single-cell resolution, we performed a mixed species experiment using the mixture of human cells (HEK) and mouse cells (3T3). After bead-cell pairing, we utilized fluorescent imaging to confirm 156 human cells and 80 mouse cells successfully paired with barcoded beads without two cells from different species in the same chamber (Figure 4-7). In our design, the cell captured in a chamber prevents the next cell being captured in the same chamber. Even if a cell tailgates after another cell into the chamber, the second cell will be removed during washing. Using shallow sequencing

Figure 4-8 Single CTC isolation and barcoded bead pairing process. (a) The cell loading process highlights the enriched CTC samples with residual blood cells and debris in the chambers. (b) The sample cleaning step removes the background cells and debris to achieve contamination-free RNA-sequencing. (c) The barcoded beads are introduced to the chamber to pair with single cells. (Scale bar: 65µm)

(~60,000 reads for cell), the mixed species sequencing results demonstrated highly organism-specific libraries without any mixed genotype cell (Figure 4-7). Using 800 genes

per cell as a threshold, we successfully recovered 147 human and 84 mouse cells, consistent with the number of cells observed under fluorescent microscopy. In addition, the HEK cells processed on different chips showed similar gene expression profile. The mixed species experiment result highlights low loss of cells, zero cellular cross contamination, and good reproducibility as important attributes of the system, distinguishing Hydro-Seq from droplet-based single-cell transcriptome sequencing technologies.

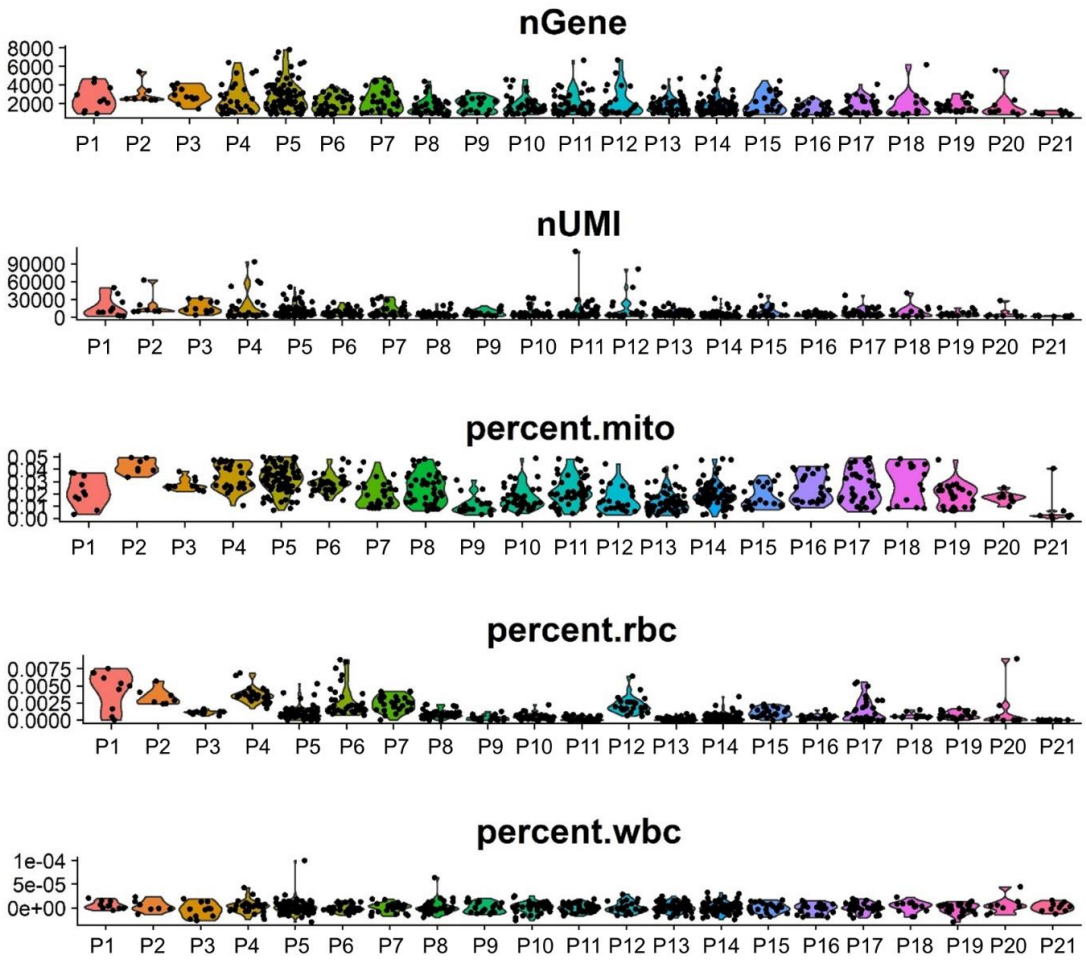

Figure 4-9 The number of genes (nGENE), the number of transcripts (nUMI), the percentage of mitochondrial genes (percent.mito), the percentage of hemoglobin genes (percent.rbc), and the percentage of CD45 gene (percent.wbc) of each patient sample. Each dot represents a CTC, and each column represents a patient sample

4.8 Circulating tumor cells sequencing

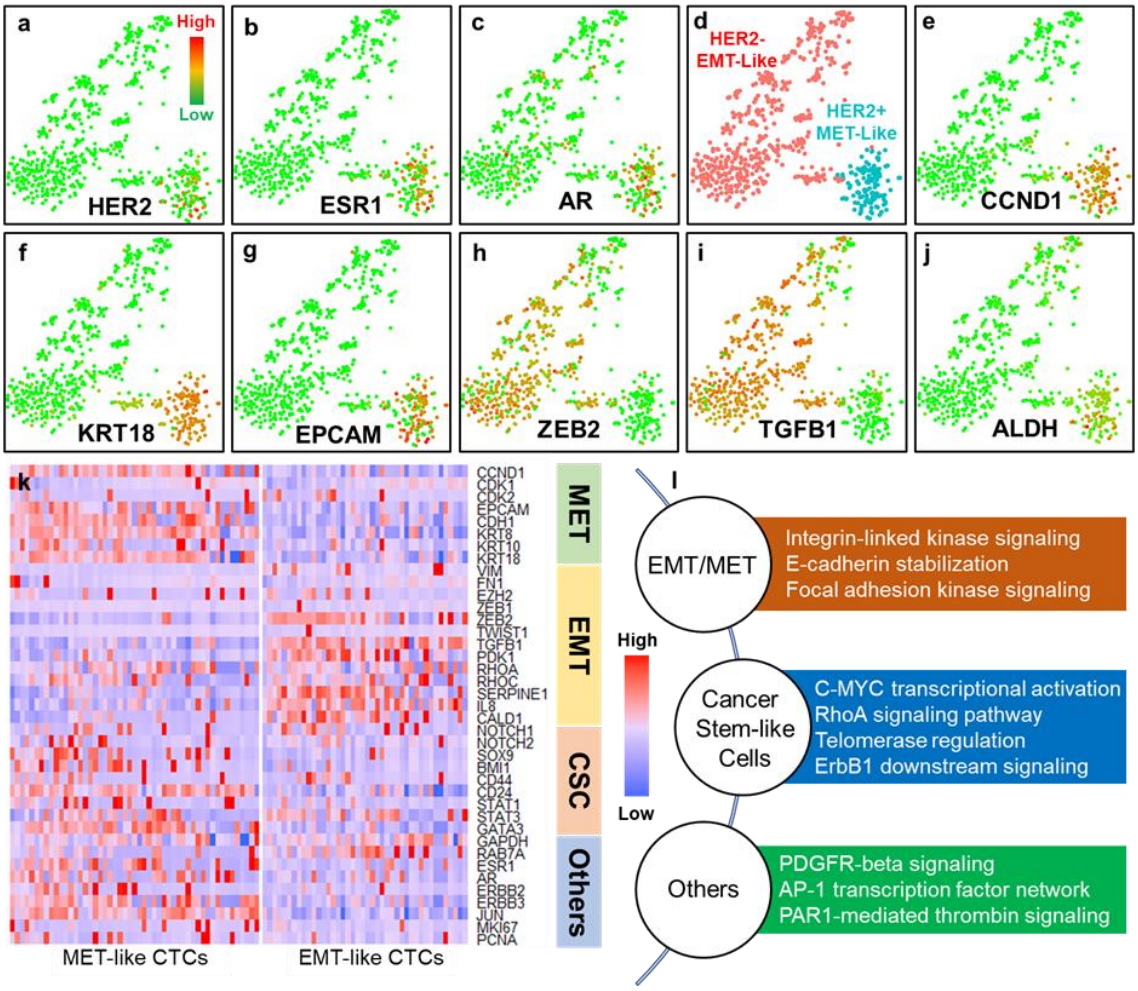

Figure 4-10 Gene expression, clustering, and pathway analysis of breast CTCs. (a-c) The expression of critical clinical markers: (a) human epidermal growth factor receptor 2 (HER2/Erbb2), (b) estrogen receptor (ESR1), and (c) androgen receptor (AR). Each dot represents one CTC. Green color represents the lowest expression, and red color represents the highest expression. The expression is logarithmically normalized. 666 CTCs from 21 patient samples were plotted based on tSNE clustering method. (d) The clustering and separation of HER2+ MET-Like and HER2- EMT-Like CTCs. (e-j) The expression of EMT and MET genes: (e) cyclin D1 (CCND1), (f) Keratin-18 (KRT18), (g) Epithelial Cell Adhesion Molecule (EPCAM), (h) EMT transcription factor (ZEB2), (i) transforming growth factor β (TGFB1), and (j) PanALDH genes (ALDH). (k) Heatmap showing the separation of HER2+ MET-Like and HER2- EMT-Like CTCs from patient P5. (l) Top-ranked pathways distinguishing the HER2+ MET-Like and HER2- EMT-Like CTC populations.

We utilized Hydro-Seq to perform single-cell transcriptomic analysis of 666 CTC's obtained from 21 metastatic breast cancer patient samples enriched by Celsee as well as by

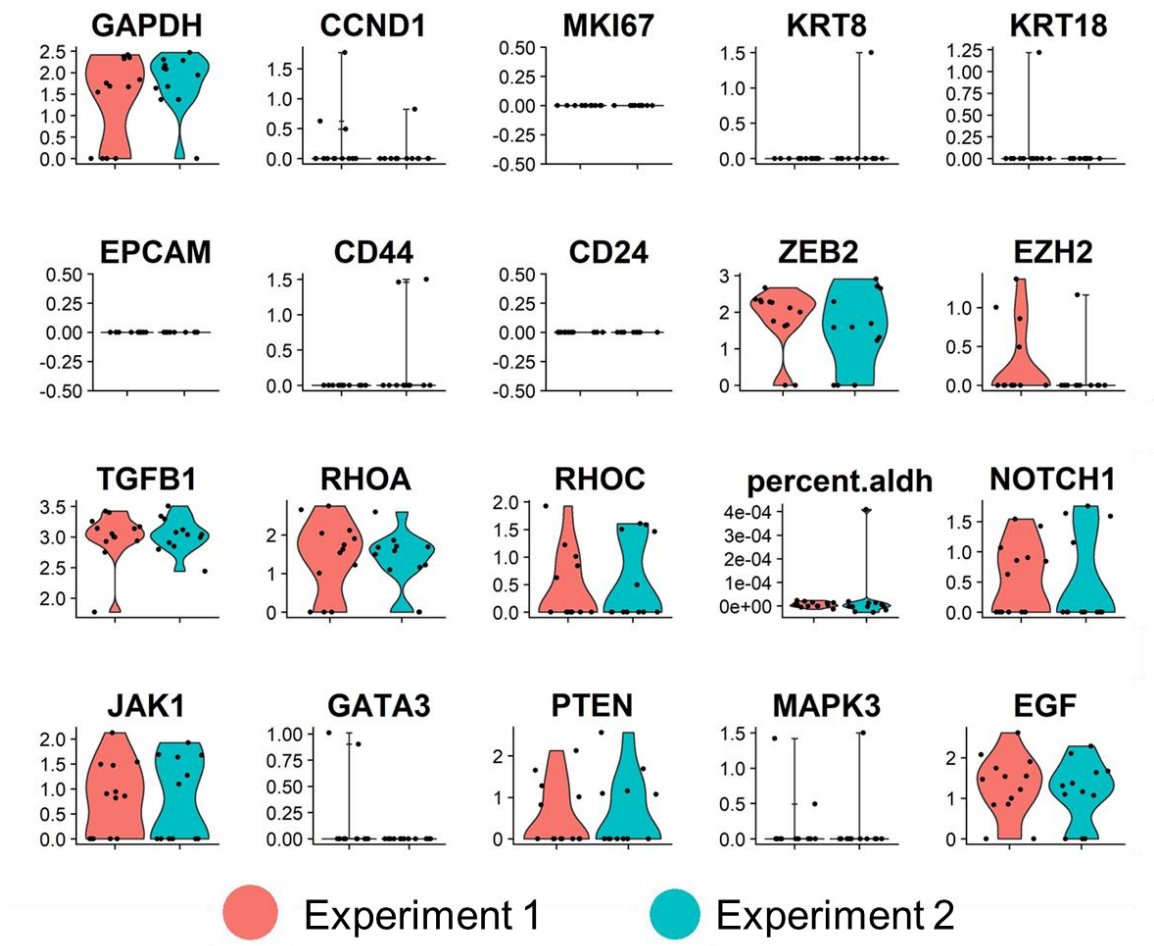

Figure 4-11 The reproducibility test of Hydro-Seq processing. For the two tubes of blood drawn from the same patient and processed on the same day, comparable CTC counts were achieved from the two experiments (Exp. 1: 13 CTCs, Exp. 2: 12 CTCs). The expression profiles of housekeeping, cell proliferation, epithelial, mesenchymal, and other genes are consistent, showing good reproducibility.

previously described microfluidic technology termed Labyrinth [6]. Captured CTCs and leukocytes could be stained on-chip to verify cell identification (Figure 4-4). After sequencing, contaminating cells were removed based on the expression of CD45 (PTPRC) and hemoglobin (Figure 4-9). The low percentage of leukocytes and erythrocytes detected, attesting the ability of Hydro-Seq to eliminate these contaminating cell populations. Gene expression data were analyzed utilizing R-based Seurat kit. To demonstrate good reliability of our approach, we processed two tubes of blood drawn from the same patient on the same

day. We detected comparable number of CTCs (Exp. 1: 13 CTCs, Exp. 2: 12 CTCs) and a similar CTC gene expression profile from two tubes (Supplementary Fig.8). We were able to detect expression of clinically important breast cancer bio-markers such as human

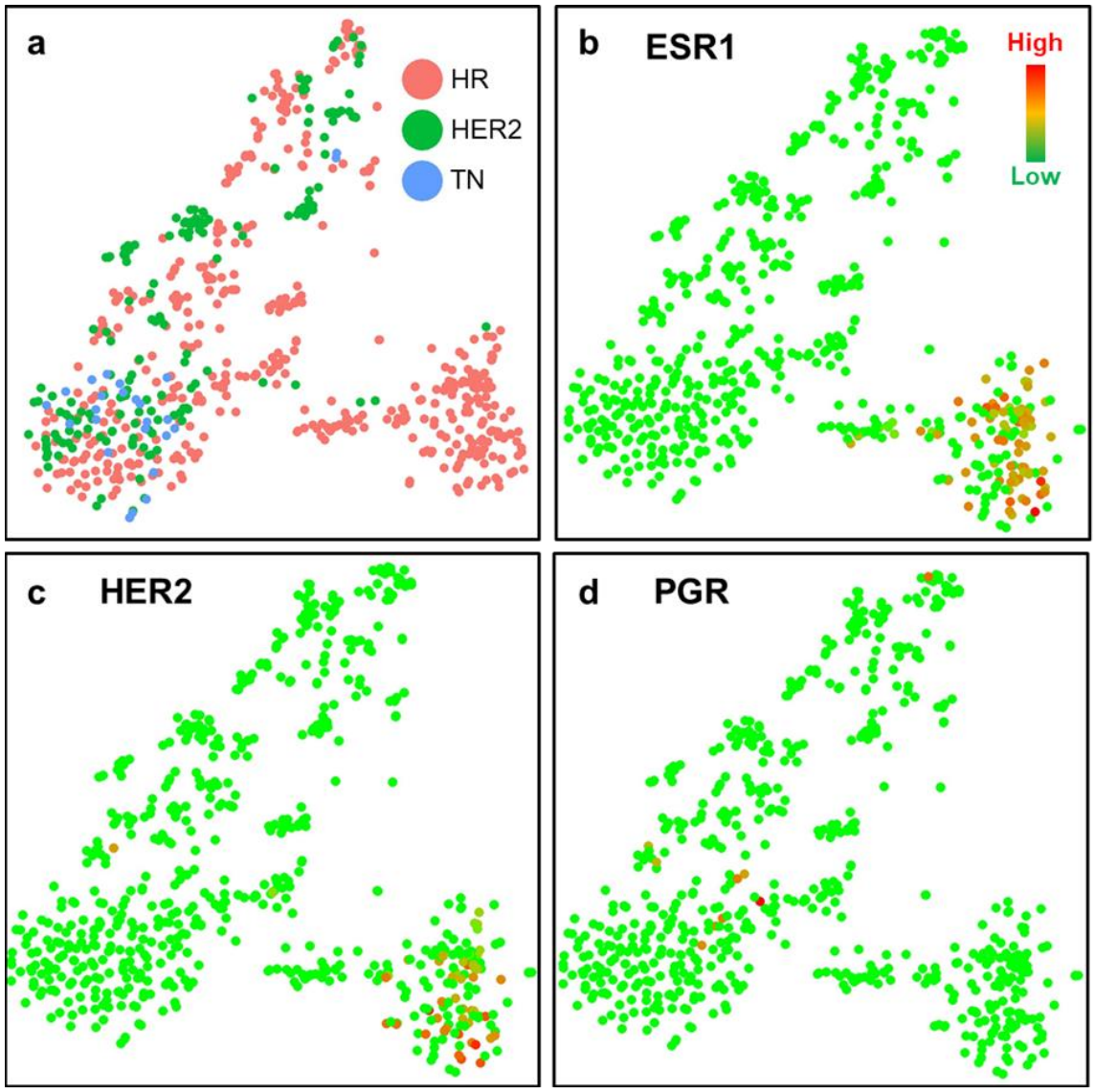

Figure 4-12 The comparison of CTCs between different molecular subtypes of primary tumor. (a) The tSNE plot of 666 CTCs from 21 patient samples. Red dot represents a CTC from hormone (estrogen or progesterone) receptor (HR) positive patient, green dot represents a CTC from HER2 positive patient, and blue dot represents a CTC from triple negative (TN) patient. (b-d) The expression of critical clinical markers: (b) estrogen receptor (ESR1), (c) human epidermal growth factor receptor 2 (HER2/Erbb2), and (d) progesterone receptor (PGR). Each dot represents one CTC. Green color represents the lowest expression, and red color represents the highest expression. The expression is logarithmically normalized.

epidermal growth factor receptor 2 (HER2/Erbb2), Estrogen receptor (ESR1), as well as Androgen receptor (AR) (Figure 4-10) [10], validating the value of cancer liquid biopsy using Hydro-Seq. Consistent with previous reports, we observed discordant molecular profiles in CTCs compared to the reported phenotype of primary tumors [11].

When further investigating the molecular characteristics of CTCs, we found they were readily separated by tSNE to two groups: HER2+ mesenchymal-to-epithelial transition (MET)-like and HER2- epithelial-to-mesenchymal transition (EMT)-like CTCs [12, 13]. The separation is independent of the number of genes, the number of unique molecules, and the percentage of mitochondrial genes. HER2+ MET-like CTCs have been previously reported to be proliferative, consistent with high expression of cyclin D1 (CCND1) [14]. In addition, MET-like CTCs express epithelial markers, including Epithelial Cell Adhesion Molecule (EPCAM) and Keratin-18 (KRT18), while HER2- EMT-like CTCs show EMT transcription factor ZEB2 and transforming growth factor β (TGFB1) [15-17]. These results confirm and extend previous studies demonstrating of the heterogeneous nature of CTCs [18]. In addition to EMT/MET, we further examined the expression of Pan-ALDH genes which are expressed in stem-like cancer cells. As expected, ALDH isoforms are over-expressed in the MET-like CTCs. Notably, we identified a CTC, expressing remarkably high Aldehyde dehydrogenase (ALDH1A3) and CD90 (THY1) locating between EMT and MET sub-populations [19-20]. This supports recent studies suggesting that CTC simultaneously expressing epithelial and mesenchymal markers may be endowed with the highest metastatic potential. This unique cell would not be detected by conventional low-efficiency low-throughput CTC profiling methods.

Though CTCs from one patient are generally cluster together, there are considerable intra-patient CTC heterogeneity. Using our single-cell resolution, we highlight a patient sample having both HER2+ MET-like and HER2- EMT-like CTCs (Fig.3k) [18]. This intra-patient heterogeneity would be easily missed by pooling CTCs from one patient altogether. Though EMT/MET cells can be identified by staining markers, with whole transcriptome sequencing, we could further investigate the difference in pathway regulation between subpopulations, identifying the activation of ITG linked kinase signaling, E-cadherin and Adherens junctions, and FAK kinases involved in EMT/MET state transitions (Figure 4-10l). Considerable studies have demonstrated that tumor metastasis is dependent on the plasticity of cancer stem-like cells (CSCs) to transition between epithelial and mesenchymal states. Consistent with this concept, we found that the EMT and MET CTCs have distinguishable activation of crucial CSC regulation pathways including c-Myc, PDGFR, RhoA, telomerase and ERB1 signaling. These results

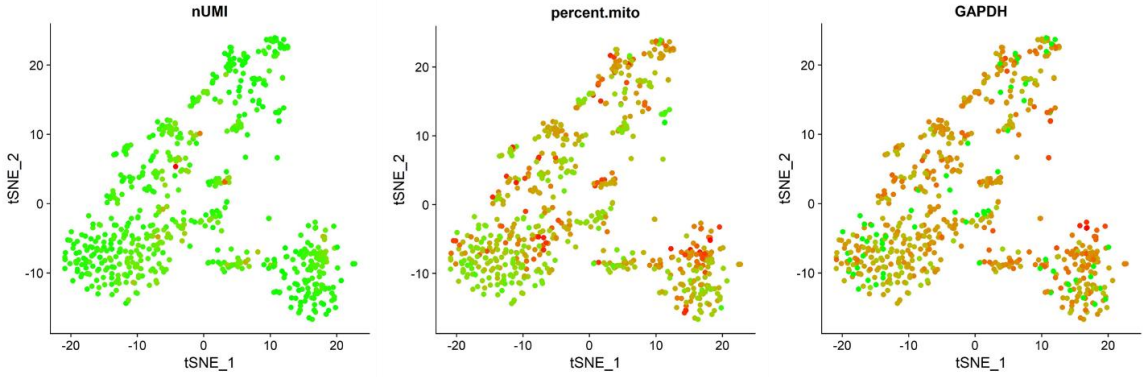

Figure 4-13 The number of transcripts (UMI), percentage of mitochondrial genes (percent.mito), and the expression of housekeeping gene GAPDH detected per cell. Each dot represents one CTC. Green color represents the lowest expression, and red color represents the highest expression. The expression is logarithmically normalized. 666 CTCs from 21 patient samples were plotted based on tSNE clustering method. The plot suggests that the separation of clusters is not determined by number of transcripts and cell viability.

demonstrate the potential of Hydro-Seq and downstream transcriptomic analysis to provide insights into the biology of CTCs and cancer metastasis.

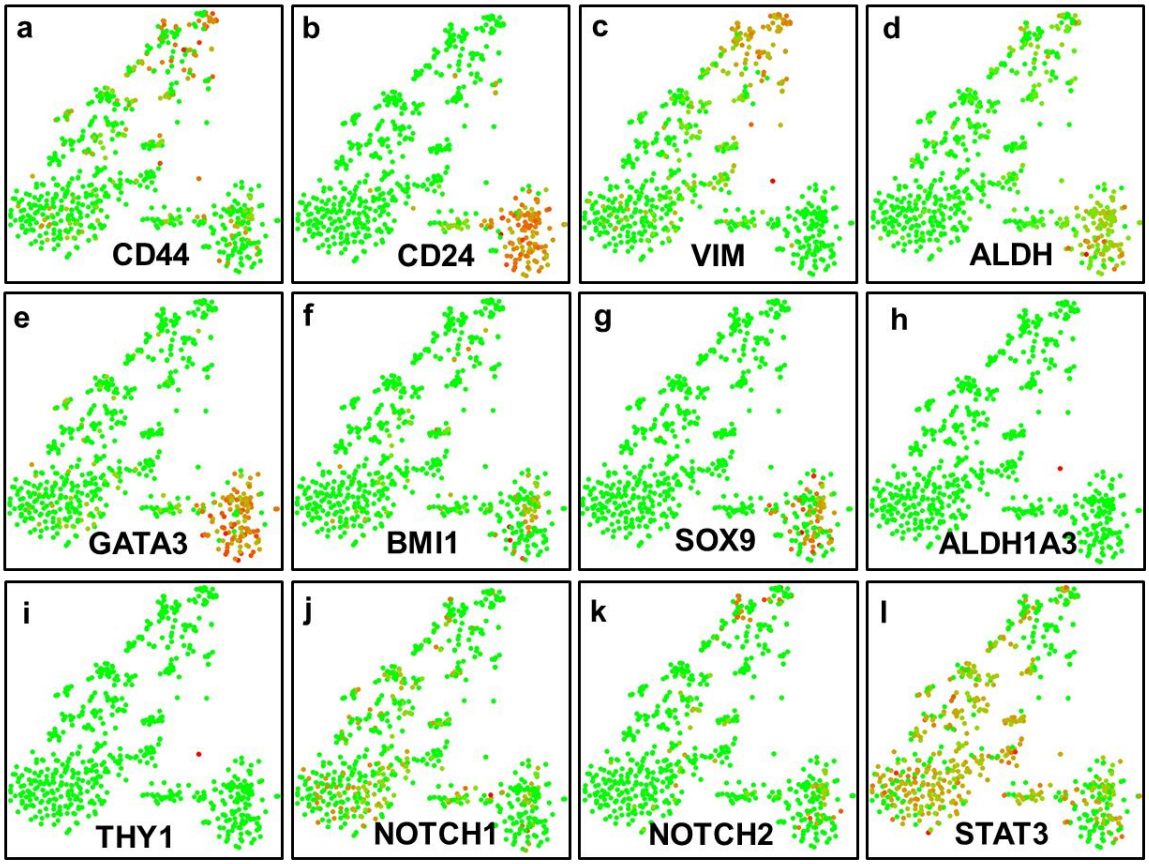

Figure 4-14 Gene expression of breast CTCs. (a-b) The expression of cancer-stem-like cell (CSC) markers: (a) CD44, a mesenchymal CSC marker. (b) CD24, an epithelial cell marker. (c) Vimentin (VIM), a mesenchymal gene, overlaps well with the CD44+ /CD24- population. (d) Pan-ALDH (same with Fig.3j). (e) GATA3 transcription factor (GATA3). (f) Polycomb complex protein BMI-1 (BMI1). (g) Transcription factor SOX-9 (SOX9). (h) Aldehyde dehydrogenase (ALDH1A3). (i) CD90 (THY1). (j) Notch homolog 1 (NOTCH1). (k) Notch homolog 2 (NOTCH2). (l) Signal transducer and activator of transcription 3 (STAT3). Each dot represents one CTC. Green color represents the lowest expression, and red color represents the highest expression. The expression is logarithmically normalized. 666 CTCs from 21 patient samples were plotted based on tSNE clustering method.

Table 4-1 The patient molecular subtypes of each sample.

Sample	Patient Subtypes	Number of CTCs	CTC Enrichment Methods
P1	HR+	10	Labyrinth
P2	HR+	7	Labyrinth
P3	HR+	11	Labyrinth
P4	HR+	29	Labyrinth
P5	HR+	78	Labyrinth
P6	HER2+	36	Labyrinth
P7	HR+	29	Labyrinth
P8	HR+	40	Labyrinth
P9	HER2+	20	Celsee
P10	HER2+	42	Celsee
P11	HR+	51	Celsee
P12	HR+	24	Celsee
P13	HR+	65	Celsee
P14	HR+	83	Celsee
P15	HR+	24	Celsee
P16	TN	25	Celsee
P17	HER2+	35	Celsee
P18	HER2+	14	Celsee
P19	HER2+	25	Celsee
P20	HR+	8	Celsee
P21	HR+	10	Celsee

4.9 Conclusion

In conclusion, we have developed Hydro-Seq platform and successfully performed scRNA-seq for 666 CTCs obtained from 21 patient samples. With the capability of

analyzing rare cell populations such as CTCs, Hydro-Seq represents a significant advance over existing technologies providing: (1) Size-based CTC isolation eliminates the biases inherent in marker-based isolation, enabling analysis of heterogeneous CTC sub-populations. (2) High-efficiency and high-throughput CTC profiling facilitates the discovery and molecular characterization of rare CTC subtypes. (3) Whole transcriptome sequencing of CTCs provides a comprehensive transcriptomic analysis compared to the much more limited capacity of immunofluorescence assays. The ability to perform efficient, comprehensive transcriptomic analysis of CTCs should greatly facilitate the use of liquid biopsy in obtaining biologically and clinically relevant information including treatment selection and therapeutic monitoring.

Chapter 5
CONCLUSION AND FUTURE WORKS

This thesis presents microfluidic-based single-cell analysis technologies for interfacing rare cell populations. Three single-cell phenotypic and transcriptomic assays are integrated on-chip to test different hypotheses in the field of cancer biology. Such single-cell analysis deciphers cellular heterogeneity of rare cell populations to facilitate cancer research, ultimately enabling clinical applications for diagnostics and treatment selection.

5.1 Contributions

5.1.1 Novel single-cell capture schemes with high single-cell capture efficiency

- *A scalable high-efficiency single cell capture scheme (chapter 2):* This highly parallel cell capture technology can be scaled from 800 chambers/chip up to 12,800 chambers/chip with optimal capture rate around 76%. With simple gravity loading, the device can be disseminated to biology labs without the need of specialized microfluidic loading equipment such as vacuum pumps or syringe pumps. With the capability to capture thousands of cells per chip, high-throughput studies can be achieved to identify rare subpopulations such as cancer stem-like cells. This is an ideal capture scheme for rare cells mixed in a large number of background cells (CSCs vs. non-CSCs), so that the rare cells can be distinguished and identified after on-chip assays.

- *A vacuum-driven single-cell capture scheme for rare samples (1-100 cells) in a small volume (1-5μL) (chapter 3):* As the vacuum force drives all the sample volume into the cell capture chamber, dead volume is minimized to prevent cell loss, achieving high single cell capture efficiency with Poisson distribution. This is an ideal single-cell capture technique for high-purity samples with a few micro-liters volume, such as the cells harvested from other microfluidics and cells harvested by cell pickers.
- *High-efficiency single cell capture scheme with contamination removal capability (chapter 4, or Hydro-Seq)*: The scalable single cell capture scheme achieves a high cell capture efficiency around 90% for rare input cells (1-100 cells). By adding on-chip valving and an additional channel (bead capture channel) to each chamber, we can perform on-chip washing to remove the debris and other background cells remaining in the chamber. This design is ideal for handling contaminated rare samples such as the circulating tumor cells. Since the sample often still contains many erythrocytes and leukocytes in the background after enrichment, Hydro-Seq can capture those rare CTCs and remove the background contamination for contamination-free downstream analysis.

5.1.2 Software development for high-throughput image analysis

- *Automatic image analysis for assay readout (Chapter 2, or μFAST)*: With thousands of cells captured on-chip, there is a strong need to develop a computer algorithm that can analyze the large area microscopy images to acquire the assay results quickly and accurately. μFAST can measure the number of cells in each chamber with ~99% accuracy for both adherent and suspension cells to calculate

capture rate and capture efficiency. For single-cell-derived sphere formation assays, μFAST is capable of quantifying the size of single cells and spheres to identify the cancer stem-like cells. Furthermore, μFAST can be extended to measure different parameters and perform more functions for different microfluidics assays.

5.1.3 Microfluidic phenotypic and transcriptomic assay

- *Single-cell-derived sphere assay (Chapter 2):* Tumor initiating cells (TICs), or cancer stem-like cells (CSCs), presents a rare cell population (1-5%) responsible for tumor relapse and metastasis. CSCs can be identified by their unique behavior of single-cells forming tumor spheres under suspension culture environment. With the scalable capture scheme, the high-throughput capability offers sufficient cells (>1000 cells) for sphere formation culture to reliably study the rare subpopulation of CSCs. By coating the triblock copolymer, F-108, on PDMS surface, the polyethylene oxide (PEO) group prevents the cells from adhering to the substrate, creating a suspension culture environment after cell capture. To demonstrate the single-cell derived assay, we tested a controversial hypothesis: whether larger or smaller cancer cells are more stem-like? The result suggested different correlations between sphere formation rate and initial cell size from different cell lines, indicating heterogeneity in pathway regulation among different cancer cell lines.
- *Single-cell proteolytic activity assay (Chapter 3):* Proteolytic degradation of extracellular matrix is a key process during cancer metastasis, and research has shown strong cellular heterogeneity in the cancer invasion process. To investigate the proteolytic activity heterogeneity, an enzymatic protease sensing assay was

integrated to the vacuum-driven cell loading chip. With cells captured in microfluidics, the nano-liter size chamber enables chemical sensing at single-cell level. By diffusion transportation of the FRET-based protease sensitive substrates and air isolation between chambers, single-cell proteolytic assay was applied to measure the intra- and inter- clonal heterogeneity and time dynamics of protease activities.

- *Single-cell RNA-sequencing of circulating tumor cells (CTCs) (Chapter 4, or Hydro-Seq)*: To better treat cancer, targeted therapy has been applied to act on specific molecular targets associated with cancer to eradicate cancer cells. To profile the potential targets of the cancer cells, molecular testing for CTCs represents a promising strategy to access the tumor molecular information for treatment selection. With single-cell analysis, rare tumor initiating cells can be identified and targeted during the therapy. To achieve scRNA-seq of CTCs, I integrated the barcoded bead technology to the high-efficiency single cell capture scheme for high efficiency bead cell pairing. With precise microfluidic flow control, the contaminating cells and debris can be removed after sample cleaning step. By applying Hydro-Seq to metastatic breast cancer patients, we successfully achieved whole transcriptome sequencing of 555 CTCs to identify key biomarkers of tumor metastasis and treatment selection. With whole transcriptome profiling, pathway analysis can be achieved to study the pathway regulation in different cancer subpopulations. The presented result highlights its great utility in therapeutic selection and monitoring of cancer patients.

5.2 Future works

5.2.1 Tumor microenvironment and cell-to-cell interaction study

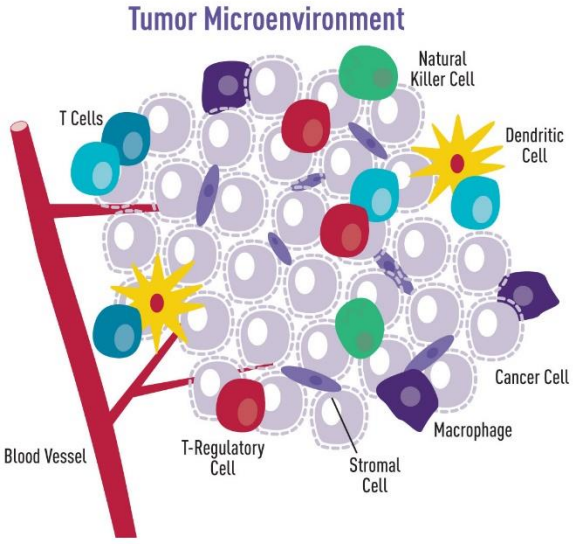

Figure 5-1 Diagram highlighting different cell types in the tumor microenvironment. Stromal cells and other immune cells interact with the tumor cell during tumor progression. [16]

Research has shown that tumor cells interact with their surrounding immune cells, fibroblasts, and other stromal cells to create a favorable environment for tumor growth and metastasis [57,58]. By studying cell-to-cell interaction in the tumor microenvironments, researchers can better understand the mechanism how tumor cells evade the immune system and recruit other stromal cells to support tumor progression. By understanding the underlying principles, better cancer treatments, such as the cancer immunotherapy, can be developed to eradicate cancer cells. To study cell-to-cell interaction, we can modify the high-throughput single cell capture chip presented in Chapter 2 to two capture sites per chamber. By supplying a mixture of tumor cells and a type of stromal cells in the cell suspension, we can pair two different cells in the same chamber and study their interaction. For example, we are interested in observing a rare cell interaction event, which some cancer

cells can engulf the neighboring immune cells to change its property. To investigate the pathway regulation related to this cell behavior, certain genes of the cancer cells can be knocked-out and tested with the engulfing assay. The capability to study cell-to-cell interaction at high-throughput will facilitate the discovery of rare yet important cell-to-cell interaction events, eventually supporting drug developments targeting regulations related to cell interactions in tumor microenvironments.

5.2.2 Cancer stem cell hierarchy study

In the cancer stem-like cell hypothesis, a tumor cell hierarchy was proposed to illustrate the process of CSCs performing self-renewal and producing downstream progenitor cells [17,59,60]. To better characterize how each cell type in the hierarchy performs self-renewal, differentiation, or de-differentiation, it is critical to observe cell division at large scale and track the molecular property of the parent cells and daughter cells. Using the single-cell capture scaling scheme presented in Chapter 2, we can achieve adherent culture of thousands of cells in different chambers. With fluorescent proteins (reporter cells) or immunohistochemistry (IHC) staining of cells, the μFAST image processing software can be applied to track the symmetric and asymmetric division events (defined by one or several biomarkers) to elucidate the cancer hierarchy structure.

After observing the cell division, we can further retrieve the daughter cells for downstream analysis. To extract the cell of interests from the chip after cell culture, a technique to selectively detach and harvest a cell from the chip was developed in the Yoon lab [41]. With the substrate coated with PDMS film on a thin layer of metal (20nm Au/Pd alloy) or carbon nanotube (6μm CNT), a nano-sec pulsed laser can be used to deform the PDMS film, creating a high temporal shear stress to detach the cells nearby. This capability

enables the downstream analysis of single cells cultured on-chip. With the cell capture scheme presented in chapter 2, the selective retrieval technology can be integrated together due to the compatibility of fabrication process and retrieval process. Utilizing PDMS-to-PDMS bonding, we can first make the capture structure with PDMS using soft-lithography and later bond it to the PDMS-metal coated substrate to create the chips for selective cell retrieval. During the retrieval operation, the parallel capture scheme allows the detached cells to flow back to inlet by simply reversing the flow from outlet to inlet. As such, selected cells can be studied using a wide variety of downstream analysis. By applying the retrieval scheme to the behavior assays, researchers can selectively retrieve cells of interests (cells from asymmetric division events or engulfing vs. non-engulfing cells) to further investigate their differences in epigenetics, transcriptomics, and proteomics activity using other measurement systems.

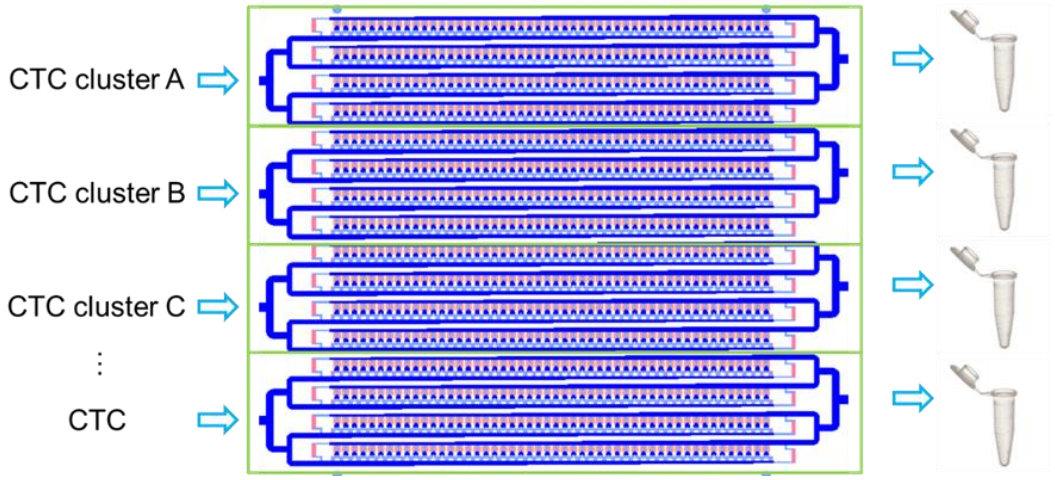

Figure 5-2 To fully investigate the gene regulation during cancer metastasis, CTC clusters should be dissociated to single cells and loaded on chip for transcriptome sequencing. Intra-cluster heterogeneity will unveil the cell composition of the cluster supporting metastasis. Comparison of transcriptome profile between single CTCs and single cells in clusters will reveal the potential genes that facilitates CTC cluster formation for cancer metastasis.

5.2.3 CTC scRNA-seq for cancer metastasis research

Tumor metastasis leads to more than 90% of cancer related death [57,61]. During the process, tumor cells break away from the primary tumor and enter the bloodstream or lymphatic system, so they can travel to a distant tissue location and potentially form a secondary or metastatic tumor [9,10,50]. To survive the hostile environment during metastasis, research has suggested that tumor cells perform state transition such as the epithelial-to-mesenchymal transition (EMT) and its reverse mesenchymal-to-epithelial transition (MET) to facilitate the migration process [14,50,57,62]. The EMT and MET cell types were observed in the metastatic cancer patient samples processed by Hydro-Seq. To further study the molecular transition during metastasis, we can apply Hydro-Seq to perform whole transcriptome analysis of tumor cells collected from the primary tumor, circulatory system, and the secondary tumor to compare their whole transcriptome profiles. In addition, research has also shown that circulating tumor cell clusters (CTC clusters) present a significantly higher metastatic potential rather than single circulating tumor cells. These circulating tumor cell clusters can be dissociated into single cells to study their whole transcriptome profile. First, this can identify the cell composition such as tumor cells, immune cells and other stromal cells in the CTC clusters, elucidating the microenvironment supporting metastasis. Second, by comparing the transcriptome profile between tumor cells in CTC clusters and single CTCs, researchers can identify the key pathway regulation and biomarkers supporting CTC cluster formation. Such thorough study of different tumor cells in tumor metastasis will unveil the upregulated and downregulated genes and pathways for the metastatic populations, and those molecular biomarkers can be eventually used for new drug development to prevent cancer metastasis, the major cause of death by cancer.

5.2.4 Clinical trials using Hydro-Seq as a companion diagnostic tool

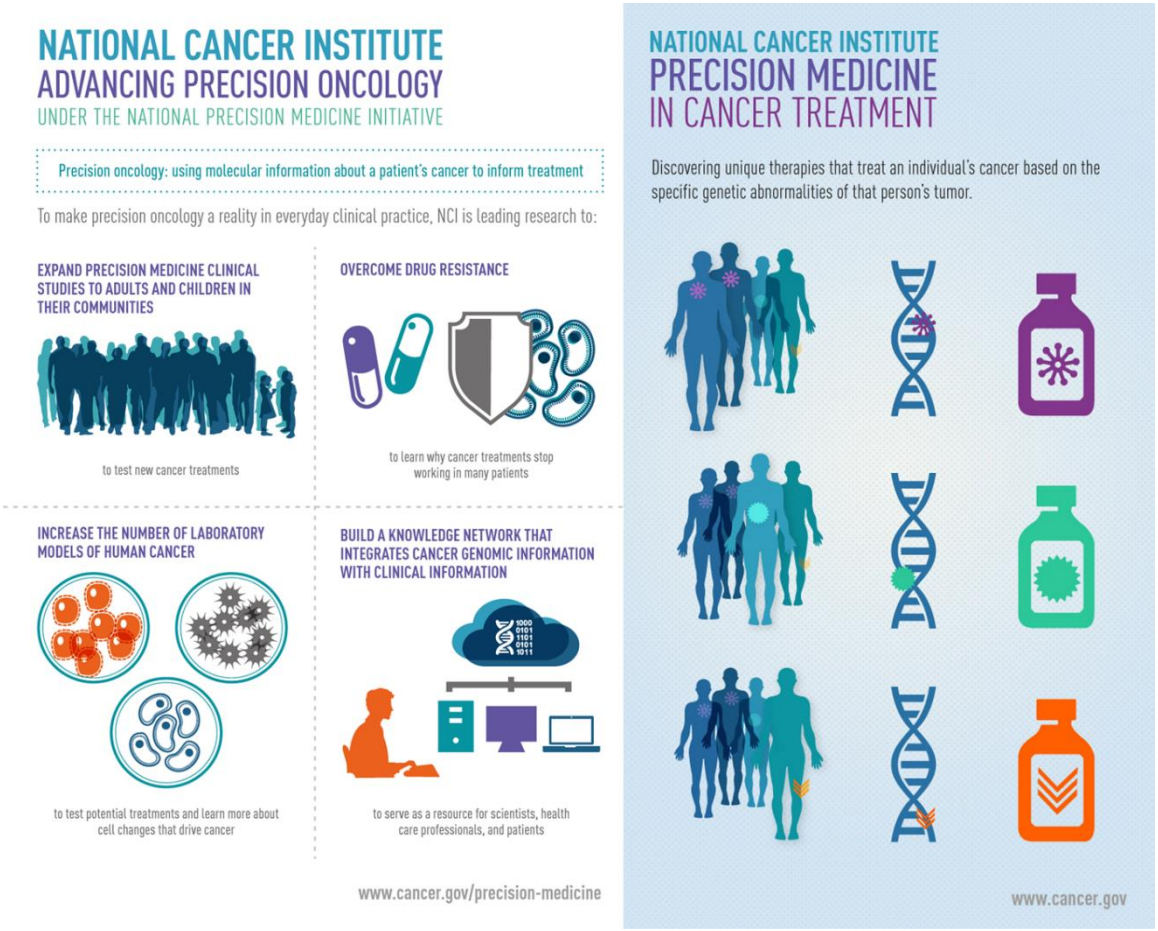

Figure 5-3 Precision medicine for cancer treatment. With the national "precision medicine initiative", the National Cancer Institute is leading research to utilize the molecular information from each patient for tailored treatment selection. One of the research direction is to expand more clinical trials and studies to identify more druggable tumor marker. Hydro-Seq presents a promising method for marker identification with simple blood draws in those clinical studies. Those clinical trials will make it possible to position Hydro-Seq as a companion diagnostic tool to inform clinicians the treatment suitable for each patient. [17] [18]

During cancer development, tumor cells have gone through many molecular and pathway regulation changes to acquire the "cancer cell" characteristics [57]. By understanding the underlying molecular characteristics of cancer cells, effective targeted therapies can be designed to target those malfunctioning molecules and pathways [63]. Due to the high degree of complexity in cancer molecular modulation, there is an emerging need

to assay cancer cells using whole transcriptome sequencing at scale, which can greatly accelerate oncology research and treatment development [64,65]. To enable such tumor transcriptome readout, Hydro-Seq can achieve whole transcriptome profiling with minimally invasive blood draws instead of the traditional highly invasive tumor biopsy, providing a strong potential for future tumor diagnostics. To achieve diagnostic applications, Hydro-Seq should be incorporated as a tool for molecular readout in early stage clinical trials with pharma company collaborators. For instance, Dr. Max Wicha, the clinician principle investigator of the Hydro-Seq project, has initiated a phase I trial with Genentech utilizing the IL6R antibody Tocilizumab to treat refractory Her2+ breast cancer.

5.2.5 Hydro-Seq chip automation and downstream integration

To disseminate the application of Hydro-Seq, it is critical to further develop Hydro-seq to be a user-friendly platform that can be easily accessed in regular research labs and hospital labs. Currently, Hydro-Seq has been applied in a lab setting with microscope, manual pipetting, syringe pumps, and a pneumatic valve control system. Those peripheral components must be integrated with the core microfluidic technology for dissemination. The reagents and valves can be controlled using the gas pressure systems (high pressure and vacuum) with all reagent preloaded into specified locations. To automate the reagent process without the need of trained technicians, the protocol can be programed into a computer system. With the chip control system, Hydro-Seq chips can be a disposable cassette to be inserted into the machine for use. With such development, standalone

systems with consumables can be disseminated to labs and other potential customers for use.

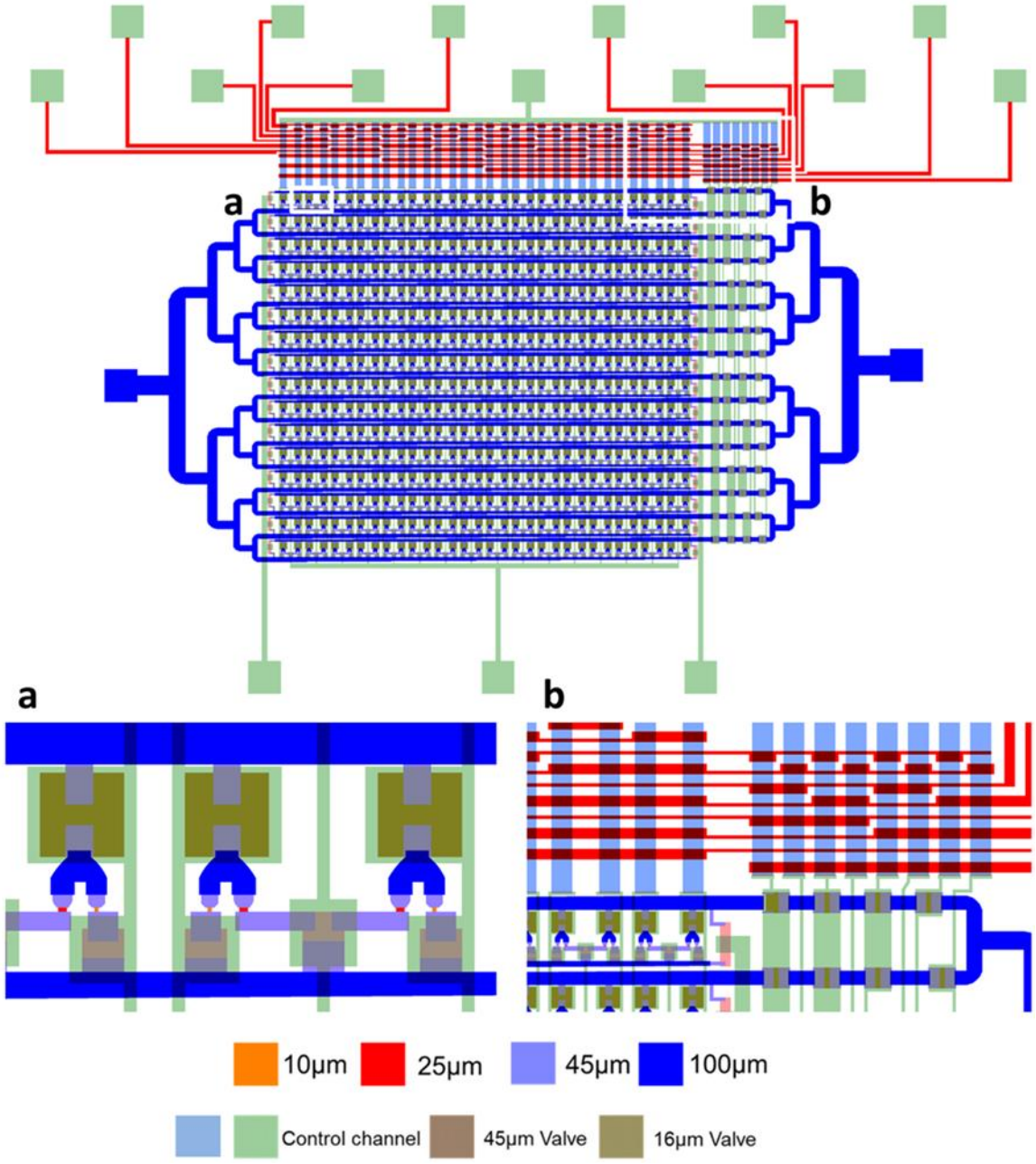

Figure 5-4 On chip multiplexer can be incorporated in Hydro-Seq to enable valve control of each chamber in Hydro-Seq. Such capability will enable selective reagent loading for specified chambers and selective retrieval of cells and beads for downstream analysis. One potential application of the chip is on-culture and analysis of cancer stem cell differentiation process.

In addition to device automation, another potential challenge for technology dissemination is the need to manually prepare the sample in wet labs for next generation sequencing. The preparation includes reverse transcription, exonuclease I treatment, and PCR amplification. Those steps can be integrated on-chip for seamless sample processing. To enable on-chip reverse transcription, researchers can implement washing chambers for the beads with microfluidic filter structures to retain beads. The addition and removal of reagents can be achieved by flowing the reagents in and out the chips. With on-chip operation, these steps can be precisely controlled and monitored to prevent losing beads during the downstream processing. For on-chip PCR, we can utilize the techniques for on-chip PCR reported in previous literature. With a flat-bed thermal cycler, reagents are loaded to the chip for PCR amplification. With the downstream sample preparation integrated on-chip, we can potentially achieve faster and higher yield reaction with less reagent usage and seamless operation.

BIBLIOGRAPHY

Chapter 1

1 A. Jemal, M. M. Center, C. DeSantis and E. M. Ward, *Cancer Epidemiol. Biomarkers & Prev.*, 2010, **19**, 1893 LP-1907.

2 C. D. Mathers, G. A. Stevens, T. Boerma, R. A. White and M. I. Tobias, *Lancet*, 2015, **385**, 540–548.

3 K. D. Miller, R. L. Siegel, C. C. Lin, A. B. Mariotto, J. L. Kramer, J. H. Rowland, K. D. Stein, R. Alteri and A. Jemal, *CA. Cancer J. Clin.*, 2016, **66**, 271–289.

4 C. Vanchieri, *JNCI J. Natl. Cancer Inst.*, 2007, **99**, 342–345.

5 F. S. Collins and H. Varmus, *N. Engl. J. Med.*, 2015, **372**, 793–795.

6 D. S. Singer, T. Jacks and E. Jaffee, *Science (80-.).*, 2016, **353**, 1105 LP-1106.

7 V. Almendro, A. Marusyk and K. Polyak, *Annu. Rev. Pathol. Mech. Dis.*, 2012, 8, 121023133009008.

8 F. A. E. Kruyt and J. J. Schuringa, *Biochem. Pharmacol.*, 2010, **80**, 423–430.

9 C. Alix-Panabières and K. Pantel, *Clin. Chem.*, 2013, **59**, 110–118.

10 M. M. Ferreira, V. C. Ramani and S. S. Jeffrey, *Mol. Oncol.*, 2016, **10**, 374–394.

11 R. Twombly, *J. Natl. Cancer Inst.* , 2005, **97**, 330–331.

12 M. D. Brooks, M. L. Burness and M. S. Wicha, *Cell Stem Cell*, 2016, **17**, 260–271.

13 C. E. Meacham and S. J. Morrison, *Nature*, 2013, **501**, 328–337.

14 T. Brabletz, *Cancer Cell*, 2016, **22**, 699–701.

15 S. Liu, Y. Cong, D. Wang, Y. Sun, L. Deng, Y. Liu, R. Martin-Trevino, L. Shang, S. P. McDermott, M. D. Landis, S. Hong, A. Adams, R. D'Angelo, C. Ginestier, E. Charafe-Jauffret, S. G. Clouthier, D. Birnbaum, S. T. Wong, M. Zhan, J. C. Chang and M. S. Wicha, *Stem Cell Reports*, 2014, **2**, 78–91.

16 B. Aktas, M. Tewes, T. Fehm, S. Hauch, R. Kimmig and S. Kasimir-Bauer, *Breast*

Cancer Res., 2009, **11**, R46.

17 T. W. Owens and M. J. Naylor, *Front. Physiol.*, 2013, **4 AUG**, 1–10.

18 J. E. Visvader and G. J. Lindeman, *Nat Rev Cancer*, , DOI:10.1038/nrc2499.

19 G. Farnie, F. Sotgia, M. P. Lisanti and F. S. M. P. L. Gillian Farnie, *Oncotarget*, 2015, **6**, 30472–30486.

20 A. Chapman, L. FernandezdelAma, J. Ferguson, J. Kamarashev, C. Wellbrock and A. Hurlstone, *Cell Rep.*, 2013.

21 A. Saadatpour, S. Lai, G. Guo and G.-C. Yuan, *Trends Genet.*, 2016, **31**, 576–586.

22 Y.-C. Chen, P. Ingram, Y. Luan and E. Yoon, eds. F.-G. Tseng and S. T. Santra, Springer Berlin Heidelberg, Berlin, Heidelberg, 2016, pp. 1–29.

23 Y.-C. Chen, P. Ingram, X. Lou and E. Yoon, in *MicroTAS 12*, Okinawa, Japan, 2012, pp. 106–108.

24 Y.-C. Chen, P. Ingram, X. Lou and E. Yoon, in *International Conference on Miniaturized Systems for Chemistry and Life Sciences (MicroTAS '12)*, 2012, pp. 1241–1244.

25 E. Shapiro, T. Biezuner and S. Linnarsson, *Nat. Rev. Genet.*, 2013, **14**, 618–30.

26 M. A. Miller, L. Barkal, K. Jeng, A. Herrlich, M. Moss, L. G. Griffith and D. A. Lauffenburger, *Integr. Biol. (Camb).*, 2011, **3**, 422–438.

27 F. Tang, C. Barbacioru, Y. Wang, E. Nordman, C. Lee and N. Xu, *Nat Methods.*, , DOI:10.1038/nmeth.1315.

28 D. Ramsköld, S. Luo, Y.-C. Wang, R. Li, Q. Deng, O. R. Faridani, G. A. Daniels, I. Khrebtukova, J. F. Loring, L. C. Laurent, G. P. Schroth and R. Sandberg, *Nat. Biotechnol.*, 2012, **30**, 777–82.

29 J.-P. Frimat, M. Becker, Y.-Y. Chiang, U. Marggraf, D. Janasek, J. G. Hengstler, J. Franzke and J. West, *Lab Chip*, 2011, **11**, 231–237.

30 Y.-H. Cheng, Y.-C. Chen, P. Ingram and E. Yoon, in *17th International Conference on Miniaturized Systems for Chemistry and Life SciencesSystems for Chemistry and Life Sciences*, Freiburg, Germany, 2013, pp. 350–352.

31 C.-H. Lin, Y.-H. Hsiao, H.-C. Chang, C.-F. Yeh, C.-K. He, E. M. Salm, C. Chen, I.-M. Chiu and C.-H. Hsu, *Lab Chip*, 2015, **15**, 2928–2938.

32 C. Ma, R. Fan, H. Ahmad, Q. Shi, B. Comin-Anduix, T. Chodon, R. C. Koya, C.-C. Liu, G. A. Kwong, C. G. Radu, A. Ribas and J. R. Heath, *Nat. Med.*, 2011, **17**, 738–743.

33 Y. Yang, Y. Mao, K.-S. Shin, C. O. Chui and P.-Y. Chiou, *Sci. Rep.*, 2016, **6**, 22630.

34 C. Yi, C.-W. Li, S. Ji and M. Yang, *Anal. Chim. Acta*, 2006, **560**, 1–23.

35 J. S. Marcus, W. F. Anderson and S. R. Quake, *Anal Chem.*, , DOI:10.1021/ac0513865.

36 A. K. White, K. A. Heyries, C. Doolin, M. Vaninsberghe and C. L. Hansen, *Anal. Chem.*, 2013, **85**, 7182–7190.

37 D. R. Reyes, D. Iossifidis, P.-A. Auroux and A. Manz, *Anal. Chem.*, 2002, **74**, 2623–2636.

38 Y.-H. Cheng, Y.-C. Chen, R. Brien and E. Yoon, *Lab Chip*, 2016, **16**, 3708–17.

39 T. M. Gierahn, M. H. Wadsworth, T. K. Hughes, B. D. Bryson, A. Butler, R. Satija, S. Fortune, J. C. Love and A. K. Shalek, *Nat. Methods*, 2017, **14**, 395–398.

40 Y.-C. Chen, N. I. Patrick, S. Fouladdel, S. P. McDermott, E. Azizi, M. S. Wicha and E. Yoon, *Sci. Rep.*, , DOI:10.1038/srep27301.

41 Y.-C. Chen, H. W. Baac, K.-T. Lee, S. Fouladdel, K. Teichert, J. G. Ok, Y.-H. Cheng, P. N. Ingram, A. J. Hart, E. Azizi, L. J. Guo, M. S. Wicha and E. Yoon, *ACS Nano*, , DOI:10.1021/acsnano.7b00413.

42 Y.-C. Chen, Y.-H. Cheng, P. Ingram and E. Yoon, in *20th International Conference on Miniaturized Systems for Chemistry and Life Sciences, MicroTAS 2016*, 2016.

43 G. Dontu and M. S. Wicha, *J. Mammary Gland Biol. Neoplasia*, 2005, **10**, 75–86.

44 H. Clevers, *Nat. Med.*, 2011, **17**, 313–319.

45 S. G. Allen, Y.-C. Chen, J. M. Madden, C. L. Fournier, M. A. Altemus, A. B. Hiziroglu, Y.-H. Cheng, Z. F. Wu, L. Bao, J. A. Yates, E. Yoon and S. D. Merajver, *Sci. Rep.*, , DOI:10.1038/srep39190.

46 Y.-C. Chen, B. Humphries, R. Brien, A. E. Gibbons, Y.-T. Chen, T. Qyli, H. R. Haley, M. E. Pirone, B. Chiang, A. Xiao, Y.-H. Cheng, Y. Luan, Z. Zhang, J. Cong, K. E. Luker, G. D. Luker and E. Yoon, *Sci. Rep.*, , DOI:10.1038/s41598-017-18610-5.

47 E. Lin, L. Rivera-Báez, S. Fouladdel, H. J. Yoon, S. Guthrie, J. Wieger, Y. Deol, E. Keller, V. Sahai, D. M. Simeone, M. L. Burness, E. Azizi, M. S. Wicha and S. Nagrath, *Cell Syst.*, 2017, **5**, 295–304.e4.

48 Y. Dong, A. M. Skelley, K. D. Merdek, K. M. Sprott, C. Jiang, W. E. Pierceall, J. Lin, M. Stocum, W. P. Carney and D. A. Smirnov, *J Mol Diagn*, 2013, **15**, 149–157.

49 X. Zhe, M. L. Cher and R. D. Bonfil, *Am. J. Cancer Res.*, 2011, **1**, 740–751.

50 P. Paterlini-Brechot and N. L. Benali, *Cancer Lett.*, 2007, **253**, 180–204.

51 D. Ramskold, S. Luo, Y. C. Wang, R. Li, Q. Deng and O. R. Faridani, *Nat Biotechnol.*, , DOI:10.1038/nbt.2282.

52 R. Riahi, P. Gogoi, S. Sepehri, Y. Zhou, K. Handique, J. Godsey and Y. Wang, *Int. J. Oncol.*, 2014, **45**, 1870–1878.

53 D. T. Miyamoto, Y. Zheng, B. S. Wittner, R. J. Lee, H. Zhu, K. T. Broderick, R. Desai, D. B. Fox, B. W. Brannigan, J. Trautwein, K. S. Arora, N. Desai, D. M. Dahl, L. V. Sequist, M. R. Smith, R. Kapur, C. L. Wu, T. Shioda, S. Ramaswamy, D. T. Ting, M. Toner, S. Maheswaran and D. A. Haber, *Science (80-.).*, 2015, **349**, 1351–1356.

54 A. Jakabova, Z. Bielcikova, E. Pospisilova, R. Matkowski, B. Szynglarewicz, U. Staszek-Szewczyk, M. Zemanova, L. Petruzelka, P. Eliasova, K. Kolostova and V. Bobek, *Breast Cancer Res. Treat.*, 2017, **166**, 695–700.

55 J. G. Lohr, V. A. Adalsteinsson, K. Cibulskis, A. D. Choudhury, M. Rosenberg, P. Cruz-Gordillo, J. M. Francis, C. Z. Zhang, A. K. Shalek, R. Satija, J. J. Trombetta, D. Lu, N. Tallapragada, N. Tahirova, S. Kim, B. Blumenstiel, C. Sougnez, A. Lowe, B. Wong, D. Auclair, Ef. M. Van Allen, M. Nakabayashi, R. T. Lis, G. S. M. Lee, T. Li, M. S. Chabot, A. Ly, M. E. Taplin, T. E. Clancy, M. Loda, A. Regev, M. Meyerson, W. C. Hahn, P. W. Kantoff, T. R. Golub, G. Getz, J. S. Boehm and J. C. Love, *Nat. Biotechnol.*, 2014, **32**, 479–484.

56 https://en.wikipedia.org/wiki/File:CSC_vs_stochastic.pdf

57 https://en.wikipedia.org/wiki/File:Linear_vs_branched.pdf

58 https://upload.wikimedia.org/wikipedia/commons/archive/3/32/20120123094828%21Fluorescence_Assisted_Cell_Sorting_%28FACS%29_B.jpg

59 https://commons.wikimedia.org/wiki/File:Fluorescence_Assisted_Cell_Sorting_(FACS)_B.jpg

60 https://en.wikipedia.org/wiki/Cancer_stem_cell#/media/File:Cancer_stem_cells_text_resized.svg

61 https://commons.wikimedia.org/wiki/File:EMT.png

62 https://en.wikipedia.org/wiki/Metastasis#/media/File:Metastasis_illustration.jpg

63 https://commons.wikimedia.org/wiki/File:CtDNA_in_circulation.png

Chapter 2

1 H. Clevers, *Nat. Med.*, 2011, **17**, 313–319.

2 F. Li, B. Tiede, J. Massagué and Y. Kang, *Cell Res.*, 2007, **17**, 3–14.

3 T. Reya, S. J. Morrison, M. F. Clarke and I. L. Weissman, *Nature*, 2001, **414**, 105–111.

4 J. A. Magee, E. Piskounova and S. J. Morrison, *Cancer Cell*, 2012, 21, 283–296.

5 V. Almendro, A. Marusyk and K. Polyak, *Annu. Rev. Pathol. Mech. Dis.*, 2012, 8, 121023133009008.

6 S. Ricardo, A. F. Vieira, R. Gerhard, D. Leitao, R. Pinto, J. F. Cameselle-Teijeiro, F. Milanezi, F. Schmitt and J. Paredes, *J. Clin. Pathol.*, 2011, 64, 937–946.

7 P. Marcato, C. A. Dean, C. A. Giacomantonio and P. W. Lee, *Cell Cycle*, 2011, **10**, 1378–1384.

8 T. W. Owens and M. J. Naylor, *Front. Physiol.*, 2013, **4 AUG**, 1–10.

9 G. Dontu and M. S. Wicha, *J. Mammary Gland Biol. Neoplasia*, 2005, **10**, 75–86.

10 F. A. E. Kruyt and J. J. Schuringa, *Biochem. Pharmacol.*, 2010, **80**, 423–430.

11 L. Vermeulen, M. Todaro, F. de Sousa Mello, M. R. Sprick, K. Kemper, M. Perez Alea, D. J. Richel, G. Stassi and J. P. Medema, *Proc. Natl. Acad. Sci.* , 2008, **105** , 13427–13432.

12 J. Stingl, *J. Pathol.*, 2009, **217**, 229–241.

13 M. J. Liao, C. Z. Cheng, B. Zhou, D. B. Zimonjic, S. A. Mani, M. Kaba, A. Gifford, F. Reinhardt, N. C. Popescu, W. Guo, N. E. Elinor, H. F. Lodish and R. A. Weinberg, *Cancer Res.*, 2007, **67**, 8131–8138.

14 E. Shapiro, T. Biezuner and S. Linnarsson, *Nat. Rev. Genet.*, 2013, **14**, 618–30.

15 Q. Li, K. Rycaj, X. Chen and D. G. Tang, *Semin. Cancer Biol.*, 2015, **35**, 191–199.

16 R. N. Aravalli, M. Behnan Sahin, E. N. K. Cressman and C. J. Steer, *Biochem. Biophys. Res. Commun.*, 2010, **391**, 56–62.

17 I. Bortolomai, S. Canevari, I. Facetti, L. De Cecco, G. Castellano, A. Zacchetti, M. R. Alison and S. Miotti, *Cell Cycle*, 2010, **9**, 1194–1206.

18 G. Farnie, F. Sotgia, M. P. Lisanti and F. S. M. P. L. Gillian Farnie, *Oncotarget*, 2015, **6**, 30472–30486.

19 V. K. Srivastava and J. Nalbantoglu, *Cytom. Part A*, 2008, **73A**, 940–948.

20 S. Zhang, I. Mercado-Uribe, Z. Xing, B. Sun, J. Kuang and J. Liu, *Oncogene*, 2014, **33**, 116–128.

21 Y.-C. Chen, P. Ingram, Y. Luan and E. Yoon, in *Essentials of Single-Cell Analysis: Concepts, Applications and Future Prospects*, eds. F.-G. Tseng and S. T. Santra,

Springer Berlin Heidelberg, Berlin, Heidelberg, 2016, pp. 1–29.

22 Y.-C. Chen, P. Ingram, X. Lou and E. Yoon, in *International Conference on Miniaturized Systems for Chemistry and Life Sciences (MicroTAS '12)*, 2012, pp. 1241–1244.

23 Y.-C. Chen, N. I. Patrick, S. Fouladdel, S. P. McDermott, E. Azizi, M. S. Wicha and E. Yoon, *Sci. Rep.,* 6, 27301

24 J. Chung, Y. J. Kim and E. Yoon, *Appl. Phys. Lett.*, 2011, **98**., 123701

25 Y.-C. Chen, Y.-H. Cheng, H. S. Kim, P. N. Ingram, J. E. Nor and E. Yoon, *Lab Chip*, 2014, **14**, 2941–7.

26 J.-P. Frimat, M. Becker, Y.-Y. Chiang, U. Marggraf, D. Janasek, J. G. Hengstler, J. Franzke and J. West, *Lab Chip*, 2011, **11**, 231–237.

27 W.-H. Tan and S. Takeuchi, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 1146–1151.

28 Y.-C. Chen, P. Ingram and E. Yoon, *Analyst*, 2014, **139**, 6371–6378.

29 A. Huebner, D. Bratton, G. Whyte, M. Yang, A. J. DeMello, C. Abell and F. Hollfelder, *Lab Chip*, 2009, **9**, 692–698.

30 V. Lecault, M. Vaninsberghe, S. Sekulovic, D. J. H. F. Knapp, S. Wohrer, F. Viel, T. Mclaughlin, A. Jarandehei, M. Miller, D. Falconnet, K. Adam, D. G. Kent, M. R. Copley, F. Taghipour, C. J. Eaves, R. K. Humphries, M. James and C. L. Hansen, *Nat. Methods*, 2011, **8**, 9–11.

31 J. R. Rettig and A. Folch, *Anal. Chem.*, 2005, **77**, 5628–5634.

32 S. H. Kim, M. Antfolk, M. Kobayashi, S. Kaneda, T. Laurell and T. Fujii, *Lab Chip*, 2015, **15**, 4356–4363.

33 C.-H. Lin, Y.-H. Hsiao, H.-C. Chang, C.-F. Yeh, C.-K. He, E. M. Salm, C. Chen, I.-M. Chiu and C.-H. Hsu, *Lab Chip*, 2015, **15**, 2928–2938.

34 K. Chung, C. A. Rivet, M. L. Kemp and H. Lu, *Anal. Chem.*, 2011, **83**, 7044–7052.

35 S. Hong, Q. Pan and L. P. Lee, *Integr. Biol.*, 2012, **4**, 374.

36 V. a. Liu, W. E. Jastromb and S. N. Bhatia, *J. Biomed. Mater. Res.*, 2002, **60**, 126–134.

37 W. Hellmich, J. Regtmeier, T. T. Duong, R. Ros, D. Anselmetti and A. Ros, *Langmuir*, 2005, **21**, 7551–7557.

38 R. Hummel, *Comput. Graph. Image Process.*, 1977, **6**, 184–195.

39 S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. T. H.

Romeny and J. B. Zimmerman, *Comput. Vis. Graph. Image Process.*, 1987, **39**, 355–368.

40 T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman and A. Y. Wu, *IEEE Trans. Pattern Anal. Mach. Intell.*, 2002, **24**, 881–892.

41 C. Zhang, X. Xiao, X. Li, Y.-J. Chen, W. Zhen, J. Chang, C. Zheng and Z. Liu, *Sensors*, 2014, **14**, 16128–16147.

42 H. Yuen, J. Princen, J. Illingworth and J. Kittler, *Image Vis. Comput.*, 1990, **8**, 71–77.

Chapter 3

1 D. Hanahan and R. A. Weinberg, *Cell*, 2011, 144, 646–674.

2 T. Ludwig, *Bioessays*, 2005, **27**, 1181–1191.

3 J. E. Koblinski, M. Ahram and B. F. Sloane, *Clin. Chim. Acta*, 2000, 291, 113–135.

4 M. L. Moss and F. H. Rasmussen, *Anal. Biochem.*, 2007, **366**, 144–148.

5 M. A. Miller, L. Barkal, K. Jeng, A. Herrlich, M. Moss, L. G. Griffith and D. A. Lauffenburger, *Integr. Biol. (Camb).*, 2011, **3**, 422–438.

6 S. J. Altschuler and L. F. Wu, *Cell*, 2010, 141, 559–563.

7 J. A. Magee, E. Piskounova and S. J. Morrison, *Cancer Cell*, 2012, 21, 283–296.

8 A. Chapman, L. FernandezdelAma, J. Ferguson, J. Kamarashev, C. Wellbrock and A. Hurlstone, *Cell Rep.*, 2013.

9 Y. Shin, S. Han, E. Chung and S. Chung, *Integr. Biol. (Camb).*, 2014, **6**, 654–61.

10 J. E. Purvis and G. Lahav, *Cell*, 2013, 152, 945–956.

11 J. E. Purvis, K. W. Karhohs, C. Mock, E. Batchelor, A. Loewer and G. Lahav, *Science*, 2012, **336**, 1440–4.

12 B. Turk, *Nat. Rev. Drug Discov.*, 2006, **5**, 785–799.

13 M. J. Meyer, J. M. Fleming, M. A. Ali, M. W. Pesesky, E. Ginsburg and B. K. Vonderhaar, *Breast Cancer Res.*, 2009, **11**, R82.

14 A. M. Streets, X. Zhang, C. Cao, Y. Pang, X. Wu, L. Xiong, L. Yang, Y. Fu, L. Zhao, F. Tang and Y. Huang, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 7048–53.

15 A. K. White, M. Vaninsberghe, O. I. Petriv, M. Hamidi, D. Sikorski, M. A. Marra, J.

Piret, S. Aparicio and C. L. Hansen, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 13999–14004.

16 A. K. White, K. A. Heyries, C. Doolin, M. Vaninsberghe and C. L. Hansen, *Anal. Chem.*, 2013, **85**, 7182–7190.

17 C. M. Hosfield, J. S. Elce, P. L. Davies and Z. Jia, *EMBO J.*, 1999, **18**, 6880–6889.

18 A. J. Hughes, D. P. Spelke, Z. Xu, C.-C. Kang, D. V Schaffer and A. E. Herr, *Nat. Methods*, 2014, **11**, 749–55.

19 T. Jing, R. Ramji, M. E. Warkiani, C. T. Lim, J. Han and C.-H. Chen, in *17th International Conference on Miniaturized Systems for Chemistry and Life Sciences*, 2013, pp. 1373–1375.

20 K. J. Son, D.-S. Shin, T. Kwa, J. You, Y. Gao and A. Revzin, *Lab Chip*, 2015, -.

21 L. Wu and J. Han, in *18th International Conference on Miniaturized Systems for Chemistry and Life Sciences*, 2014, pp. 846–848.

22 W.-H. Tan and S. Takeuchi, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 1146–1151.

23 J. Nilsson, M. Evander, B. Hammarström and T. Laurell, *Anal. Chim. Acta*, 2009, 649, 141–157.

24 Y.-C. Chen, P. Ingram and E. Yoon, *Analyst*, 2014, **139**, 6371–6378.

25 C. Lochovsky, S. Yasotharan and A. Günther, *Lab Chip*, 2012, 12, 595.

26 D. M. Bickett, M. D. Green, J. Berman, M. Dezube, A. S. Howe, P. J. Brown, J. T. Roth and G. M. McGeehan, *Anal. Biochem.*, 1993, **212**, 58–64.

27 A. P. Patel, I. Tirosh, J. J. Trombetta, A. K. Shalek, S. M. Gillespie, H. Wakimoto, D. P. Cahill, B. V Nahed, W. T. Curry, R. L. Martuza, D. N. Louis, O. Rozenblatt-Rosen, M. L. Suvà, A. Regev and B. E. Bernstein, *Science*, 2014, **344**, 1396–401.

28 V. Almendro, A. Marusyk and K. Polyak, *Annu. Rev. Pathol. Mech. Dis.*, 2012, 8, 121023133009008.

29 S. Liu, Y. Cong, D. Wang, Y. Sun, L. Deng, Y. Liu, R. Martin-Trevino, L. Shang, S. P. McDermott, M. D. Landis, S. Hong, A. Adams, R. D'Angelo, C. Ginestier, E. Charafe-Jauffret, S. G. Clouthier, D. Birnbaum, S. T. Wong, M. Zhan, J. C. Chang and M. S. Wicha, *Stem Cell Reports*, 2014, **2**, 78–91.

30 Y.-C. Chen, P. Ingram, X. Lou and E. Yoon, in *International Conference on Miniaturized Systems for Chemistry and Life Sciences (MicroTAS '12)*, 2012, pp. 1241–

1244.

31 R. B. Keithley, E. M. Weaver, A. M. Rosado, M. P. Metzinger, A. B. Hummon and N. J. Dovichi, *Anal. Chem.*, 2013, **85**, 8910–8918.

32 K. R. Love, S. Bagh, J. Choi and J. C. Love, *Trends Biotechnol.*, 2013, 31, 280–286.

33 F. Li, B. Tiede, J. Massagué and Y. Kang, *Cell Res.*, 2007, **17**, 3–14.

34 P. Marcato, C. A. Dean, C. A. Giacomantonio and P. W. Lee, *Cell Cycle*, 2011, **10**, 1378–1384.

35 S. Ricardo, A. F. Vieira, R. Gerhard, D. Leitao, R. Pinto, J. F. Cameselle-Teijeiro, F. Milanezi, F. Schmitt and J. Paredes, *J. Clin. Pathol.*, 2011, 64, 937–946.

Chapter 4

1 E. Z. Macosko, A. Basu, R. Satija, J. Nemesh, K. Shekhar, M. Goldman, I. Tirosh, A. R. Bialas, N. Kamitaki, E. M. Martersteck, J. J. Trombetta, D. A. Weitz, J. R. Sanes, A. K. Shalek, A. Regev and S. A. McCarroll, *Cell*, 2016, **161**, 1202–1214.

2 T. M. Gierahn, M. H. Wadsworth, T. K. Hughes, B. D. Bryson, A. Butler, R. Satija, S. Fortune, J. C. Love and A. K. Shalek, *Nat. Methods*, 2017, **14**, 395–398.

3 A. M. Klein, L. Mazutis, I. Akartuna, N. Tallapragada, A. Veres, V. Li, L. Peshkin, D. A. Weitz and M. W. Kirschner, *Cell*, 2016, **161**, 1187–1201.

4 H. C. Fan, G. K. Fu and S. P. A. Fodor, *Science (80-.).*, 2015, **347**.

5 M. M. Ferreira, V. C. Ramani and S. S. Jeffrey, *Mol. Oncol.*, 2016, **10**, 374–394.

6 E. Lin, L. Rivera-Báez, S. Fouladdel, H. J. Yoon, S. Guthrie, J. Wieger, Y. Deol, E. Keller, V. Sahai, D. M. Simeone, M. L. Burness, E. Azizi, M. S. Wicha and S. Nagrath, *Cell Syst.*, 2017, **5**, 295–304.e4.

7 Y.-H. Cheng, Y.-C. Chen, R. Brien and E. Yoon, *Lab Chip*, 2016, **16**, 3708–17.

8 R. Riahi, P. Gogoi, S. Sepehri, Y. Zhou, K. Handique, J. Godsey and Y. Wang, *Int. J. Oncol.*, 2014, **45**, 1870–1878.

9 I. Bozovic-Spasojevic, D. Zardavas, S. Brohée, L. Ameye, D. Fumagalli, F. Ades, E. de Azambuja, Y. Bareche, M. Piccart, M. Paesmans and C. Sotiriou, *Clin. Cancer Res.*, 2017, **23**, 2702 LP-2712.

10 A. Jakabova, Z. Bielcikova, E. Pospisilova, R. Matkowski, B. Szynglarewicz, U. Staszek-Szewczyk, M. Zemanova, L. Petruzelka, P. Eliasova, K. Kolostova and V. Bobek, *Breast Cancer Res. Treat.*, 2017, **166**, 695–700.

11 S. Liu, Y. Cong, D. Wang, Y. Sun, L. Deng, Y. Liu, R. Martin-Trevino, L. Shang, S. P. McDermott, M. D. Landis, S. Hong, A. Adams, R. D'Angelo, C. Ginestier, E. Charafe-Jauffret, S. G. Clouthier, D. Birnbaum, S. T. Wong, M. Zhan, J. C. Chang and M. S. Wicha, *Stem Cell Reports*, 2014, **2**, 78–91.

12 N. V. Jordan, A. Bardia, B. S. Wittner, C. Benes, M. Ligorio, Y. Zheng, M. Yu, T. K. Sundaresan, J. A. Licausi, R. Desai, R. M. O'Keefe, R. Y. Ebright, M. Boukhali, S. Sil, M. L. Onozato, A. J. Iafrate, R. Kapur, D. Sgroi, D. T. Ting, M. Toner, S. Ramaswamy, W. Haas, S. Maheswaran and D. A. Haber, *Nature*, 2016, **537**, 102.

13 E. Shaulian and M. Karin, *Oncogene*, 2001, **20**, 2390.

14 J. H. Tsai and J. Yang, *Genes Dev.*, 2013, **27**, 2192–2206.

15 V. Karantza, *Oncogene*, 2010, **30**, 127.

16 C. Vandewalle, J. Comijn, B. De Craene, P. Vermassen, E. Bruyneel, H. Andersen, E. Tulchinsky, F. Van Roy and G. Berx, *Nucleic Acids Res.*, 2005, **33**, 6566–6578.

17 M. Yu, A. Bardia, B. S. Wittner, S. L. Stott, M. E. Smas, D. T. Ting, S. J. Isakoff, J. C. Ciciliano, M. N. Wells, A. M. Shah, K. F. Concannon, M. C. Donaldson, L. V Sequist, E. Brachtel, D. Sgroi, J. Baselga, S. Ramaswamy, M. Toner, D. A. Haber and S. Maheswaran, *Science (80-.).*, 2013, **339**, 580 LP-584.

18 M. Toloudi, P. Apostolou, M. Chatziioannou and I. Papasotiriou, *Case Rep. Oncol.*, 2011, **4**, 44–54.

19 P. Marcato, C. A. Dean, D. Pan, R. Araslanova, M. Gillis, M. Joshi, L. Helyer, L. Pan, A. Leidal, S. Gujar, C. A. Giacomantonio and P. W. K. Lee, *Stem Cells*, 2011, **29**, 32–45.

20 M. V. Shaikh, M. Kala and M. Nivsarkar, *Cancer Biomarkers*, 2016, **16**, 301–307.

Chapter 5

1 D. Hanahan and R. A. Weinberg, *Cell*, 2011, 144, 646–674.

2 Y.-C. Chen, Y.-H. Cheng, H. S. Kim, P. N. Ingram, J. E. Nor and E. Yoon, *Lab Chip*, 2014, **14**, 2941–7.

3 T. W. Owens and M. J. Naylor, *Front. Physiol.*, 2013, **4 AUG**, 1–10.

4 S. Ricardo, A. F. Vieira, R. Gerhard, D. Leitao, R. Pinto, J. F. Cameselle-Teijeiro, F. Milanezi, F. Schmitt and J. Paredes, *J. Clin. Pathol.*, 2011, 64, 937–946.

5 D. Iliopoulos, H. A. Hirsch, G. Wang and K. Struhl, *Proc. Natl. Acad. Sci. U. S. A.*,

2011, **108**, 1397–1402.

6 Y.-C. Chen, H. W. Baac, K.-T. Lee, S. Fouladdel, K. Teichert, J. G. Ok, Y.-H. Cheng, P. N. Ingram, A. J. Hart, E. Azizi, L. J. Guo, M. S. Wicha and E. Yoon, *ACS Nano*, , DOI:10.1021/acsnano.7b00413.

7 F. Li, B. Tiede, J. Massagué and Y. Kang, *Cell Res.*, 2007, **17**, 3–14.

8 P. Paterlini-Brechot and N. L. Benali, *Cancer Lett.*, 2007, **253**, 180–204.

9 C. Alix-Panabieѐres and K. Pantel, *Clin. Chem.*, 2013, **59**, 110–118.

10 M. M. Ferreira, V. C. Ramani and S. S. Jeffrey, *Mol. Oncol.*, 2016, **10**, 374–394.

11 M. J. Bissell and W. C. Hines, *Nat. Med.*, 2011, **17**, 320.

12 T. Brabletz, *Cancer Cell*, 2016, **22**, 699–701.

13 C. Sawyers, *Nature*, 2004, **432**, 294.

14 B. T. Hennessy, A. M. Gonzalez-Angulo, M. S. Carey and G. B. Mills, *Clin. Cancer Res.*, 2009, **15**, 417–419.

15 S. Kimbung, N. Loman and I. Hedenfalk, *Semin. Cancer Biol.*, 2015, **35**, 85–95.

16 https://visualsonline.cancer.gov/details.cfm?imageid=11067

17 https://visualsonline.cancer.gov/details.cfm?imageid=10171

18 https://visualsonline.cancer.gov/details.cfm?imageid=9895